Unravelling the stepwise activation mechanism of HacA, the key regulator of the Unfolded Protein Response in *Aspergillus niger*

Harm Mulder

The work described in this thesis was carried out at Danisco Genencor within framework 5 of Eurofung, and was supported by EC Grant No: QLRK3-00729. Publication of this thesis was financially supported by Danisco Genencor.

Printed by: Ipskamp Drukkers

ISBN: 978-90-9024977-3

Unravelling the stepwise activation mechanism of HacA, the key regulator of the Unfolded Protein Response in *Aspergillus niger*

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 2 februari 2010
klokke 15:00 uur
door

Harm Jan Mulder

geboren te Oosterend-Texel
in 1970

Promotion Committee

Promoter:	**Prof. Dr. C. A. M. J. J. van den Hondel**
Co-promoter:	**Dr. J. Visser**
Other Members:	**Prof. Dr. S.M. Verduyn Lunel**
	Prof. Dr. P.J.J. Hooykaas
	Prof. Dr. P.J. Punt
	Dr. A.F.J. Ram
	Prof. Dr. D. B. Archer (University of Nottingham)
	Dr. C.M.J. Sagt (DSM)

Voor Sjuttje en Kornelius Jan

Contents

Chapter 1

General Introduction

General Introduction

Fungi perform the essential role of decomposing organic matter in nature, and are indispensable in nutrient cycling and exchange. Together with the Animalia, Plantae and Protista, Fungi compose the four kingdoms of eukaryotic organisms.

Aspergillus niger is a soil dwelling filamentous fungus that is ubiquitous in soils worldwide. Being a saprophytic organism it plays an important role in the degradation and recycling of organic debris such as plant cell-wall material. The saprophytic nature of *A. niger* requires the abundant secretion of a variety enzymes to degrade and extract nutrients from complex polymeric substrates. The released sugars can either be metabolised to sustain growth, or under certain conditions be converted into organic acids like citric acid which the fungus can re-use later. Like other filamentous fungi, *A. niger* grows in the vegetative phase of its lifecycle by forming tubular shaped cells called hyphae. A hypha is usually made up of multiple cells, separated by septa. These hyphae extend at their apices, and branch subapically resulting in a network of hyphae extending in all directions. Protein secretion occurs mostly at the hyphal tip, and together with the apical extension of the hyphae, this allows the fungus to penetrate and digest solid substrates.

The Industrial Use of *Aspergillus niger*

The ability of *A. niger* to produce citric acid was discovered by James Currie in 1917, and has been exploited by the industry since 1919. The ability to secrete large amounts of enzymes into the environment has been utilised since the 1960s for the production of enzymes for the starch processing and food industry [146]. The long use of *A. niger* in the industry for production of food enzymes and citric acid has led to the GRAS status of many of the products produced by *A. niger* (U.S. Food and Drug Administration. http://www.cfsan.fda.gov/~rdb/opa-gras.html).

The progress in genetic manipulation of filamentous fungi not only led to the optimisation of the production of homologous proteins [170], but also provided possibilities for using filamentous fungi for the production of heterologous proteins of both fungal and non-fungal origin. However, whereas large amounts of homologous proteins can be obtained in industrial fermentations, the production yields for heterologous proteins of non-fungal origin are typically several orders of magnitude lower [129,167].

By dissecting the steps involved in synthesis and secretion of protein several factors influencing the secretion level have been identified. Protein synthesis can be limited at every level in the process; from the transcriptional, translational and post-translational level to the secretory and post-secretory level. Factors that can negatively influence the

final production level are: (1) mRNA stability, (2) codon usage, (3) inefficient translocation, folding and transport through the secretion pathway and (4) extracellular degradation [47]. A number of strategies addressing those limitations have been applied in order to optimise or increase heterologous protein production by filamentous fungi. Among those strategies are the use of strong homologous promoters, the use of multicopy transformants, of protease-deficient host strains, the introduction of efficient secretion signals, the use of well-secreted proteins as carriers, and the over-expression of foldases and/or chaperones [2,47]. The right combination of the above mentioned approaches together with classical strain improvement can increase the production yield for some heterologous proteins, but often not up to the levels of homologous protein production [47].

The expression of a heterologous protein often is accompanied by an upregulation of ER localised foldases. Protein secretion and in particular protein folding and maturation in the endoplasmic reticulum (ER) is therefore regarded as the major bottleneck in heterologous protein production by filamentous fungi [2,17,47,120,139]. The efforts to improve heterologous protein production in filamentous fungi focuses therefore mostly on the secretory pathway and the folding of proteins in the ER in particular.

The Secretory Pathway

Eukaryotic cells possess a secretory pathway that is made up of an elaborate endomembrane network. This network consists of several independent organelles that make up an assembly line for the production of secretory proteins and proteins with other cellular destinations such as the cell membrane. Each compartment in the secretory pathway provides a specialised environment that facilitates the various stages in protein biogenesis, modification, sorting and finally secretion into the environment.

The high capacity of the secretory system of *A. niger* and other filamentous fungi, has been exploited for decades by the industry for the production of proteins and metabolites, but despite the intense genetic research on the protein secretion mechanism in filamentous fungi, the knowledge of the fungal secretory pathway has been limited. However, the availability of the whole genome sequence of several filamentous fungi, including the industrial important species *Aspergillus oryzae* [87] and *A. niger* [126], have made formal genomics approaches to the study of protein secretion in those organisms more feasible. It is generally accepted that the secretory pathway in filamentous fungi does not differ greatly from those in yeast and higher eukaryotes, which have been studied in more detail. The differences lay mainly in the polar growth of filamentous fungi. The mycelial growth phenotype of filamentous fungi, which consist of multi-cellular tubular hyphae made up of cylindrical compartments divided by perforated septa, is not found in yeast or higher eukaryotes. Also, protein secretion in filamentous fungi is believed to occur mainly at the hyphal tip or sub-apical hyphal regions [46,182]. So besides the capacity of the protein

secretion pathway, which is much higher than in yeast, the secretory pathway of filamentous fungi stands out mostly by the directionality of the pathway due to hyphal growth.

The secretory pathway in filamentous fungi has been reviewed by several authors [18,125,149]. Secretory proteins enter the endoplasmic reticulum (ER) upon translation, and during transit through this compartment they are folded into their tertiary and quaternary structure, and undergo modification if necessary. The modifications include signal peptide processing, disulfide bridge formation, glycosylation, and phosphorylation. Correctly folded proteins are sorted into coating-protein-II vesicles (COPII) and transported from the ER to the Golgi-like structure where further modifications occur, including glycosylation and peptide processing. The mature proteins are finally packed into secretory vesicles and directed to the plasma membrane at the hyphal tip where they are secreted to the extracellular space. Terminally misfolded proteins on the other hand are removed from the ER by retro-translocation and degraded in the cytosol by the 26S proteasome in a process called ER-associated degradation (ERAD). A schematic overview of the secretion pathway in filamentous fungi is given in Figure 1, and described in more detail below.

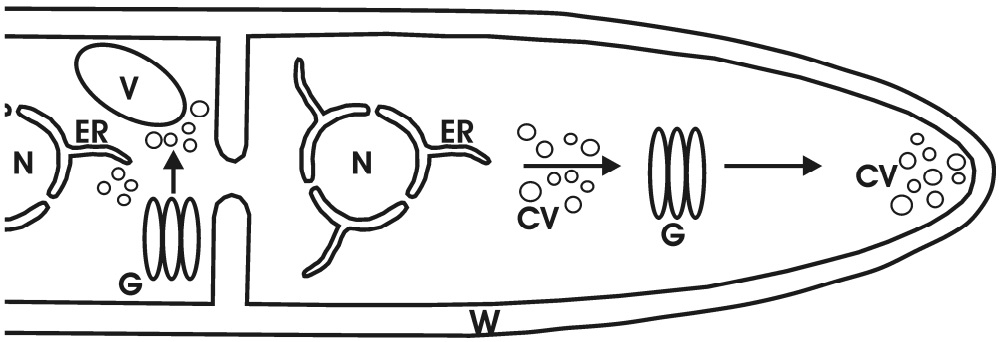


Figure 1. The secretion pathway in filamentous fungi. Newly synthesized polypeptides undergo folding and modifications in the ER. Only correctly folded proteins leave the ER (in vesicles) and enter the Golgi-like structure. There the proteins undergo further modifications and are finally secreted from the apical or subapical hyphal regions. N: Nucleus. ER: Endoplasmic reticulum. CV: Cytoplasmic vesicles. G: Golgi(-like) compartment. V: Vacuole. W: Cell wall [47].

ER Targeting and Translocation

Protein synthesis starts in the cytosol where ribosomes capture mRNA molecules, and begin translation. Proteins destined for the secretory pathway can be directed to the ER in two ways; either co-translational or post translational [131]. The co-translational pathway predominates in most mammalian cells, whereas in *S. cerevisiae* both routes are used extensively. The route that is followed is determined by the hydrophobic core of its signal sequence [103]. It was found that proteins with a less hydrophobic signal sequence were

targeted through the so called SRP-independent route, whereas both routes could be followed when a more hydrophobic signal was present.

In the co-translational translocation, an N-terminal signal sequence, which consists of 13-30 amino acids, is bound by a cytosolic signal recognition particle (SRP), as soon as it has emerged from the ribosome. The translation pauses and the complex (mRNA, nascent peptide chain, ribosome and SRP) is directed to the ER membrane via an interaction with a SRP receptor (SRPR). After docking, the SRP is released from the complex and the nascent polypeptide chain is passed through a protein translocation complex on the ER membrane upon elongation [131]. Once the polypeptide chain is completed, the signal sequence is cleaved off and the immature protein is released into the ER lumen.

Posttranslational translocation is independent of the SRP, and the translocation of proteins over the ER membrane occurs first after the polypeptide has been fully translated. This alternative route of targeting proteins to the ER membrane is thought to also occur in filamentous fungi. After synthesis in the cytoplasm, the nascent protein forms a complex with the cytosolic Hsp70 chaperone and co-chaperones to keep it in an unfolded state. The protein-chaperone complex is directed to the ER membrane by the signal sequence, where it is recognised by the Sec-complex [103] which acts as a membrane receptor. The actual translocation occurs similar as in the SRP-dependent pathway through the Sec61p channel. In *S. cerevisiae* 20 genes have been identified to play a role in the translocation process of newly synthesised proteins into the ER. Genome sequencing of *A. niger* identified 18 genes that encode yeast orthologs of those genes [126]. As in mammals, no ortholog of the *S. cerevisiae* signal recognition and docking protein Srp21p was found, a gene that is essential in yeast. Also no homologue was found for the non-essential yeast *SSH1* gene, which codes for the alpha subunit of the Ssh1 translocon complex of the ER. Despite the apparent absence of those two genes in *A. niger*, protein translocation into the ER is thought to occur analogous to yeast.

Although proteins are targeted to the ER by their signal sequence, their ultimate destination may be different. They may be destined for secretion, targeted to other cellular destinations or have the ER as their final destination. ER resident proteins are retained in the compartment due to a C-terminal tetrameric (K/HDEL) retention signal [127] which targets the protein back to the ER if it ends up in the Golgi.

Protein Folding in the Endoplasmic Reticulum (ER)

After translocation into the ER, the signal sequence of the nascent protein is removed by the signal peptidase complex [31], and the proteins undergo assisted folding and distinct modifications such as disulfide bridge formation, *cis-trans* isomerisation of peptide bonds preceding proline residues, and the initiation of glycosylation.

Proteins enter the ER in an immature form and therefore display a considerable amount of hydrophobic areas, which attract each other in the aqueous environment of the ER, and could give rise to protein aggregation. An important function of the ER is to prevent these unwanted interactions between immature proteins and to enable efficient protein folding. This is assisted by a set of ER-resident chaperones and foldases. *A. niger* encodes seventeen luminal proteins that are involved in the protein folding processes in the ER [126]. They include chaperones like BiP (binding protein) and calnexin, and foldases of the protein disulphide isomerase family (PDI) and peptidyl-prolyl *cis-trans* isomerase family (PPI).

The process of protein folding in the ER of *Aspergilli* has recently been reviewed by Geysens *et al.* [40]. Based on the analysis of the genome sequences of *A. niger* and *A. nidulans* they review various aspects protein folding in the light of knowledge from *S. cerevisiae* and *Pichia pastoris*.

Chaperones

The chaperone BiP is a member of the heatshock 70 protein family (Hsp70), and it is one of the best studied chaperones for which a homologue (*bipA*) has been cloned in *A. niger* [130,168]. BiP binds to hydrophobic stretches as soon as the growing polypeptide chain enters the ER to suppress aggregation, and it promotes the translocation of the polypeptide chain across the membrane [41]. The binding and the release of BiP to and from peptides is coupled to the hydrolysis of ATP (Fig. 2). Not only is BiP involved in binding immature proteins and thereby in the quality control of protein folding [190], it is also involved in regulating the UPR through its association with the transmembrane kinase IreA. BiP is therefore regarded as a key molecular chaperone. In addition to BipA, another putative HSP70 chaperone, LshA, was found in *A. niger*. In yeast, Lsh1p and the BipA homologue Kar2p have interactions with DnaJ proteins, Sec63 and the nucleotide exchange factor Sil1p [155]. Sil1p interacts directly with the ATPase domain of Kar2p, modulating the activity of the chaperone [163]. Surprisingly, no homologue for Sil1p was found in *A. niger* [126].

The gene encoding the *A. niger* chaperone calnexin (*clxA*) has also been cloned and characterised [17,173]. Calnexin is a lectin-like chaperone, and an essential part of the maturation and quality control mechanism of glycoproteins. It specifically interacts with partially trimmed monoglucosylated N-linked oligosaccharides. It serves as a carbohydrate-binding protein allowing for cycles of deglucosylation and reglucosylation of the maturing protein.

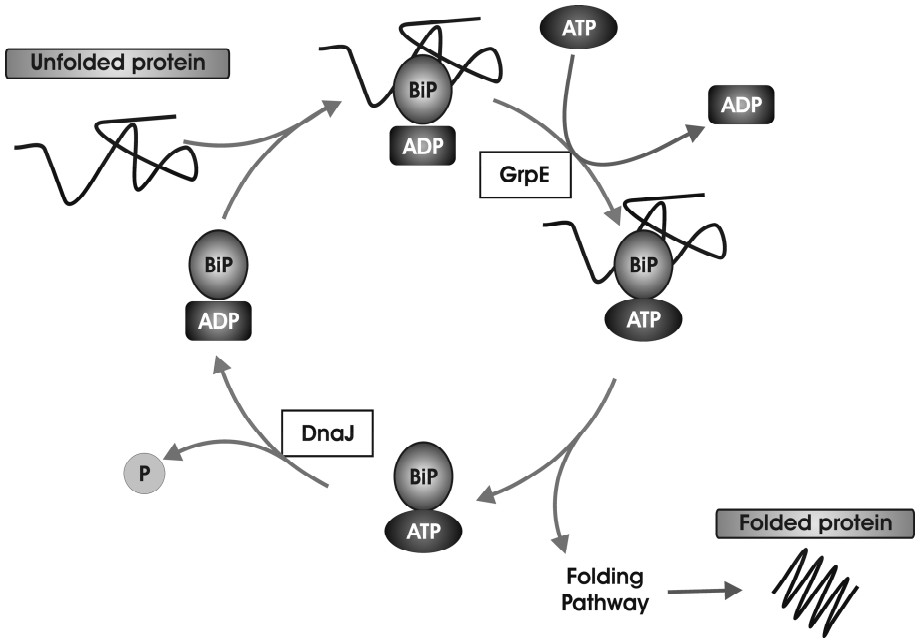


Figure 2. The BiP ADP-ATP cycle. Modified after Schröder and Kaufman [144].

Foldases

In addition to the chaperones, foldases are involved in the protein folding. Foldases catalyse slow and often rate limiting covalent changes such as the formation of disulfide bonds between cysteine residues, and the isomerisation of peptide bonds preceding proline residues.

The oxidising environment of the ER favours the formation of disulphide bonds, which are required for the stability and function of a large number of secreted proteins [32]. The formation and rearrangement of disulphide bridges is catalysed by PDI's [109]. PDI is a protein thiol-oxidoreductase that catalyses the oxidation, reduction and isomerisation of protein disulphide bonds. The genes encoding the *A. niger* PDI and several PDI-related genes have been isolated and characterised. These are *pdiA* [104] encoding PDIA, and the genes *tigA* [62] and *prpA* [174] encoding the PDI-related proteins TIGA (tunicamycin inducible gene) and PRPA (PDI-related protein). *A. niger* encodes also a putative membrane-bound PDI-family protein (EpsA) [126].

Transfer of electrons during oxidative protein folding, in for example disulphide bond formation, requires in *S. cerevisiae* the ER membrane protein Ero1p (Fig. 3) [36,162]. Homologues of this protein have also been discovered in the *Aspergillus* genomes [126].

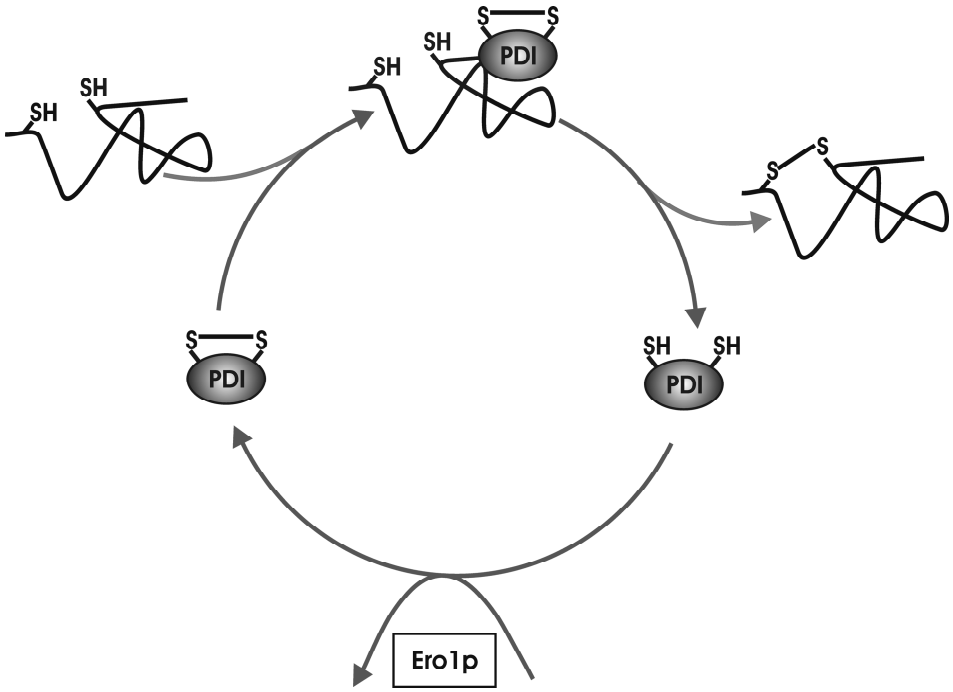


Figure 3. Oxidative protein folding mediated by PDI.

The PPIases catalyse the *cis-trans* isomerisation of peptide bonds preceding proline residues (Fig. 4). For every amino acid residue except for proline, there is steric hindrance between the neighbouring side chains for the *cis* conformation but not for the *trans* conformation. The peptide bonds for these residues are therefore all *trans*. However, in the case of proline the steric hindrance is almost equal for both isomers, hence the presence of PPIases. In native proteins more than 10% of the proline peptide bonds can be in the *cis* conformation [86]. The PPIases were discovered due to their ability to bind to immunosuppressive drugs. In filamentous fungi two major classes of PPIases have been discovered; The cyclophilins which bind to cyclosporin A (CsA), and the FK506 binding proteins (FKBPs) which can bind to the FK506 compounds. However, most of the PPIases predicted from the fungal genomes are cytosolic, and only few ER resident ones have been identified like the *cypB* gene of *A. niger* which encodes an ER resident cyclophilin-like PPI [27]. Its C-terminal HEEL ER retention signal is somewhat divergent from the K/HDEL consensus, but experimental data proved its ER localisation. The putative *A. niger* ER lumenal protein EroA also has a predicted C-terminal HEEL ER-retention signal [126]. The genome sequence of *A. niger* also revealed a gene for a second ER-localised PPIase, an FKBP-type peptidyl-prolyl *cis-trans* isomerase (An01g06670).

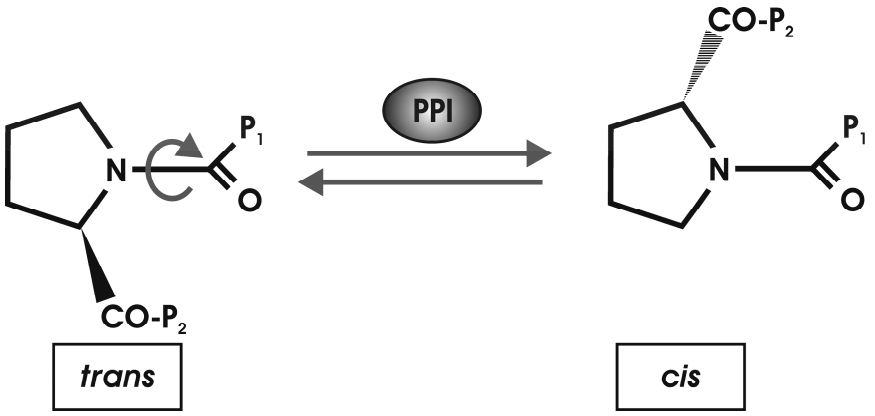


Figure 4. Isomerisation of peptide bonds preceding proline residues, catalyzed by PPIases.

Glycosylation

The majority of proteins secreted by filamentous fungi are glycosylated, a process that starts in the ER. Oligosaccharide moieties are either attached to asparagine residues in N-linked glycosylation or to serine or threonine residues in O-linked glycosylation. N-linked glycosylation of proteins, is best understood, and serves highly diverse functions. It stabilises the proteins against denaturation and proteolysis, enhances solubility, modulates immune responses, facilitates orientation of proteins relative to a membrane, confers structural rigidity to proteins, regulates protein turnover, and fine-tunes the charge and isoelectric point of proteins [55].

Glycosylation also plays an important role in the quality control mechanism of the ER [28], as it distinguishes unfolded from folded and misfolded proteins. Glycosylation starts on the nascent polypeptide chain as it enters the ER. The oligosaccharyltransferase, which is associated with the translocon, scans and glycosylates asparagine residues in the sequon N-X-S/T, where X is any amino acid except proline [55]. The oligosaccharyl-transferase adds N-linked oligosaccharides to the side chain nitrogen of the asparagine residue by an N-glycosidic bond. The N-linked glycan core in fungi has shown to be identical to the mammalian N-linked core ($Glc_3Man_9GlcNac_2$) [88]. After the initial glycosylation, glucosidase I and glucosidase II trim the oligosaccharide to the monoglucosylated form ($Glc_1Man_9GlcNac_2$), which allows the glycopeptides to interact with calnexin and calreticulin [49]. Calreticulin is the soluble counterpart of calnexin. Calreticulin is present in mammalian cells, but not in yeast and *A. niger*. Being a lectin, calnexin specifically interacts with glycoproteins, but only if they have monoglucosylated N-glycans. In mammalian cells, it has been shown that a thioredoxin family foldase (Erp57) associates with the calnexin and promotes folding [57]. Removal of the remaining glucose from the glycan by glucosidase

II results in the dissociation of the glycopeptides from calnexin. If, at this point, the glycoprotein is folded correctly, it will be no longer retained in the ER and will be transported to the Golgi complex. If not, re-addition of a glucose residue to the N-linked glycan occurs by the action of the UDPglucose:glycoprotein glucosyltransferase (UGGT), a luminal enzyme that acts as a folding sensor [134]. This folding sensor makes up the ER quality control (ERQC) system together with the lectins. The reglucosylation by UGGT tags the incomplete folded glycoprotein for renewed interaction with calnexin, and therefore a new folding attempt. The deglucosylation-reglucosylation cycle continues until proper folding is achieved (Fig. 5). The main features of the cycle are well established, and detailed aspects are reviewed in a number of papers [28,29,30,121]. Homologues were identified for nearly all components of the glycosylation pathway and *A. niger* therefore potentially possesses a well-developed glycosylation-dependent quality control system [126].

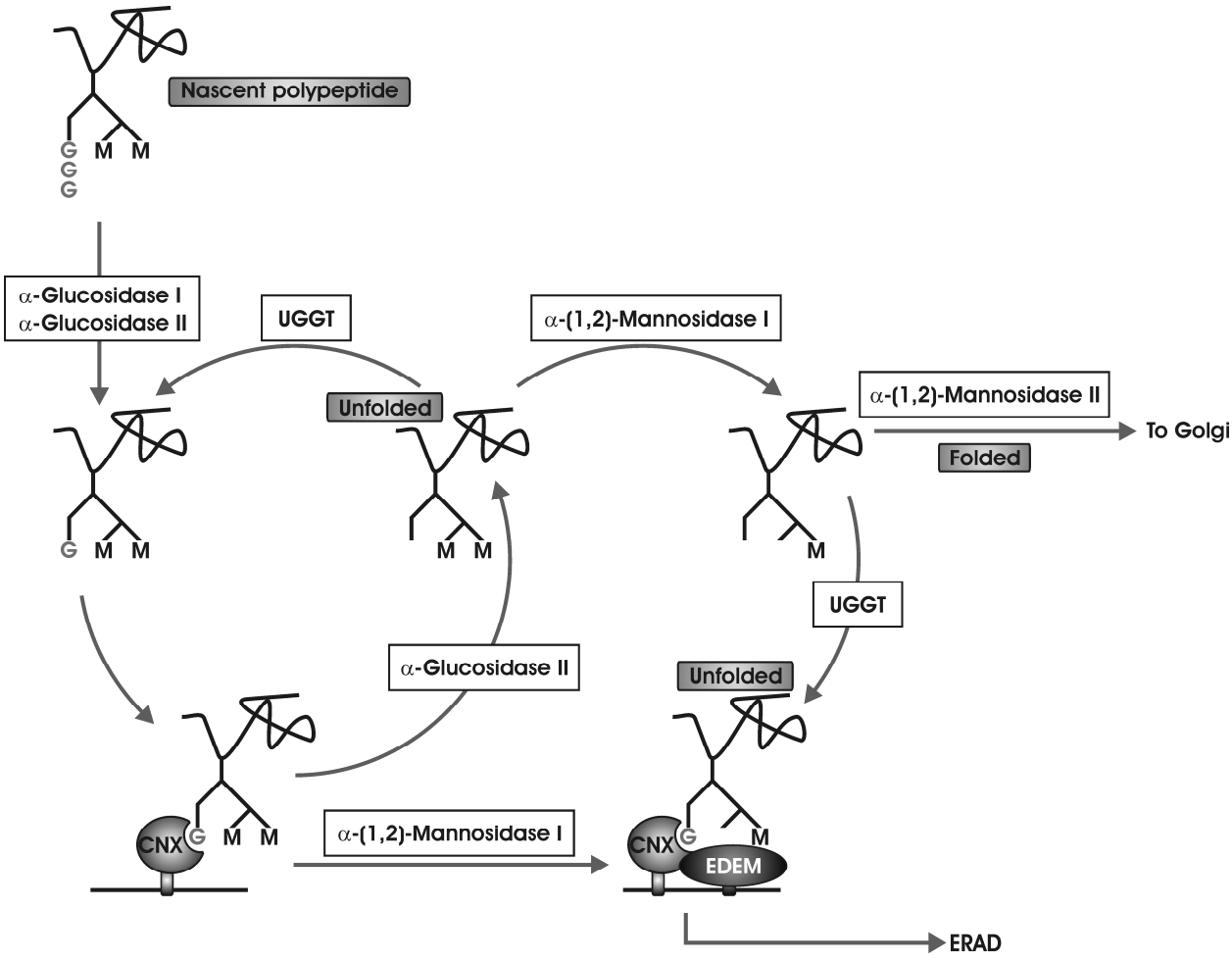


Figure 5. ER quality control mechanisms. The calnexin/calreticulin cycle. Enzymatic actions are represented in square boxes. Abbreviations: CNX, calnexin; EDEM, ER degradation-enhancing α-mannosidase-like protein; G, glucose; M, mannose; and UGGT, uridine diphosphate (UDP)-glucose:glycoprotein glucosyl transferase. Modified after Schröder and Kaufman [144].

Transport to Golgi

When folded correctly, the proteins pass the quality control system in the ER and are send off to the Golgi-like structure where they can undergo further modifications such as O-linked glycosylation, processing of N-linked oligosaccharides, phosphorylation and proteolytic processing. Although a classical Golgi-apparatus consisting of a series of flattened cisternae is not commonly seen in filamentous fungi, functions associated with the Golgi are present, and the term Golgi-like structure is normally used.

The transport of proteins between the ER and the Golgi complex, which is bidirectional, is carried out by specific vesicles. There are two types of vesicles involved in the transport between ER and Golgi. COPII vesicles (Coat Protein complex II) are involved in the anterograde transport from ER to Golgi [6], whereas COPI vesicles are involved in retrograde transport [20].

COPII

The formation of vesicles from the ER starts with the recruiting of COPII proteins to the ER membrane [82]. In yeast this is initiated by the ER-resident protein Sec12p, which serves as a guanine nucleotide exchange factor for the small GTPase Sar1p. Sar1p is active in its GTP-bound form and recruits the Sec23-Sec24 heterodimer to the ER membrane. The complex then recruits the Sec13-Sec31 complex which is thought to polymerise the coat and drive the membrane deformation to form a COPII vesicle [142]. The forming vesicle must somehow capture its protein cargo. How this is achieved is not yet fully understood, but the Sar1p-Sec23-Sec24 complex is thought to interact with cargo proteins via specific sorting signals. A universal ER export signal has not been described, instead a number of different signals seem to govern the interaction with the COPII vesicle coat. Specific transmembrane cargo proteins are sorted into the COPII coated vesicle by conserved di-acidic motifs and di-hydrophobic or aromatic motifs. The signalling motifs of these transmembrane cargo proteins are cytoplasmically exposed and recognised by subunits of the COPII coat [5]. The sorting of soluble cargo proteins from the ER is less well characterised, but there is evidence for both bulk-flow and receptor mediated sorting. The Erv29p transmembrane protein from yeast is responsible for packaging glycosylated pro-alpha-factor into COPII vesicles [7]. In addition, Erv29p is also likely to sort other soluble proteins including vacuolar hydrolases, carboxypeptidase Y, and proteinase A. However, deletion of *ERV29* suggested that additional cargo loading receptors must be present, since a number of other secretory protein were packaged normally in Erv29p depleted cells [11]. After completion, COPII vesicles travel on microtubules to the *cis*-Golgi, where they fuse with the Golgi membrane and deposit their load into the Golgi. Fusion of the COPII vesicle with the Golgi requires disassembly of the coat, which is mediated by GTP hydrolysis on Sar1p, catalysed by Sec23 [114]. The mechanism by which the transport receptors release their cargo upon fusion with

the Golgi is not all clear yet, but may rely on a change in pH or another physiological condition which triggers the dissociation of the cargo and its receptor.

COPI

COPI vesicles are involved in the retrograde transport from Golgi to ER, and shuttle mislocated ER-resident proteins back to the ER. Various ER-resident type I transmembrane proteins contain a canonical ER retrieval motif, KKXX, on their cytoplasmic domain [19]. This motif interacts with COPI coat to transport the proteins back to the ER. Soluble ER-resident proteins bear a C-terminal K/HDEL retention or retrieval signal that mediates interaction with the K/HDEL receptor, in yeast encoded by *ERD2*. Although the K/HDEL receptor does not possess a canonical cytoplasmic di-lysine retrieval motif, it is thought to also interact with the COPI coat to facilitate its transport back to the ER [10]. The precise mechanism by which the receptor binds its ligand is not known, but ligand binding is thought to trigger uptake of the assembled complex into COPI vesicles [83].

The Golgi is also important for sorting proteins to their final destinations like vacuoles or the external environment. Proteins enter the Golgi on the *cis*-side and are transported to the medial and *trans*-Golgi by vesicles, and undergo modifications meanwhile. In the *trans*-Golgi network (TGN) proteins are finally sorted to their destination. They are either transported in vesicles to the plasma membrane, where they fuse and deliver their cargo by exocytosis, or they are transported to vacuoles. Exocytosis also mediates the addition of lipids and membrane proteins to the plasma membrane. Some of the secreted proteins are required for cell wall biosynthesis, whereas others are dedicated to the hydrolysis of polymeric substrates and are released into the surroundings.

ER Stress Responses

The secretory pathway is thus responsible for the production and transport of membrane and secretory proteins and proteins with other cellular destinations like endosomes and lysosomes. The ER plays a pivotal role in the production and assembly-line of the secretory pathway, as it executes some major physiological functions. Not only are proteins synthesised at the ER, the majority of the proteins also fold into their native conformation in the ER. At the same time these nascent proteins undergo a series of post-translational modifications, such as disulphide bond formation and the addition of N-linked oligosaccharides. The ER is also an important intracellular Ca^{2+} store for the cell. Finally, it is a site where biological membranes are assembled, and it is involved in fatty acid and phospholipid biosynthesis [68,93,158]. At homeostasis, the three functions of the ER (protein folding, Ca^{2+} store and its function as a site for fatty acid and phospholipid synthesis) are in

a dynamic equilibrium. Disturbance of any of the three homeostatic functions of the ER leads to stress, and is manifested by the accumulation of unfolded proteins.

ER stress can arise transiently as a cell's gene expression program is altered in response to changing extracellular signals. For example, changing environmental conditions can result in an increased requirement for extracellular enzymes, which in turn can overwhelm the ER with newly synthesised proteins waiting to be folded. Likewise, this condition can arise in industrial stains overproducing homologous enzymes under optimised conditions. ER stress can also be caused by other environmental perturbations the cell encounters. These include nutrient starvation, anoxia, viral infections, and heat [13,60,85].

ER stress can also be more permanent in cells that bear mutations that interfere with protein maturation in the ER, or in cells that express difficult to fold heterologous proteins. Furthermore, several chemical agents can disturb the ER and lead to stress. Reducing agents such as β-mercaptoethanol or 1,4-dithiothreitol (DTT) disturb the oxidative environment of the ER and reduce disulfide bonds thereby interfering with the oxidative protein folding in the ER. The drug tunicamycin, which is produced by the bacterium *Streptomyces lysosuperificus* [157], blocks the production of N-linked glycoproteins by irreversibly inhibiting the enzyme GlcNAc phosphotransferease (GPT), resulting in accumulation of unfolded proteins in the ER. The calcium ionophore A23187 induces ER stress by depleting the ER luminal Ca^{2+} store. This in turn inhibits protein folding and interferes with multiple processes in the ER including chaperone function [117,156], UGGT activity [3], and ERAD targeting due to decreased α(1-2)-mannosidase activity [34].

To cope with and adapt to ER stress, an intracellular ER-to-nucleus signal transduction pathway has evolved in the eukaryotic cell. This signal transduction pathway is termed the Unfolded Protein Response (UPR), and activation of this pathway triggers an extensive transcriptional response which adjusts the folding capacity of the ER. The initial task of the UPR is to salvage unfolded proteins by increasing the folding capacity of the ER. However, the UPR also enhances the targeting of unfolded proteins in the ER for recycling. When proteins fail to obtain their proper fold, even after several folding attempts, they have to be removed from the ER. These misfolded proteins could otherwise accumulate, aggregate and become toxic for the cell. An efficient proteolytic system is therefore coupled to the quality control machinery of the ER to dispose of these terminally misfolded proteins.

Two routes are known that target these misfolded proteins for degradation: (i) ER-associated degradation (ERAD) in which the misfolded proteins are retrotranslocated to the cytosol, ubiquitinated and degraded by the 26S proteasome [91,135], and (ii) Autophagy in which parts of the ER are targeted to lysosomes or vacuoles [75].

The Unfolded Protein Response in *S. cerevisiae*

The UPR was first characterised in *S. cerevisiae*, where genetic screens identified three proteins that were required for the signal transduction from the ER to the nucleus: a transmembrane serine/threonine kinase, Ire1p [21], a basic leucine zipper type transcription factor Hac1p [22,95] , and a tRNA ligase, Rlg1p [151]. The role of these three components in the yeast UPR is schematised in Figure 6.

Ire1p is activated by the accumulation of unfolded proteins in the ER, and facilitates together with Rlg1p the splicing of an unconventional intron from *HAC1* mRNA. Removal of the unconventional intron from the *HAC1* mRNA enables the efficient translation of the messenger and the subsequent synthesis of active Hac1p protein. The active Hac1p translocates to the nucleus, where it promotes transcription of its target genes by an interaction with the unfolded protein response element (UPREs), present in the promoters of UPR-regulated genes such as genes encoding ER chaperones and foldases. The UPR target genes of yeast have been defined by microarray expression profiling [160]. They comprise 381 genes which is more than 5% of the yeast genome. Among the UPR target genes are many proteins that play critical roles in the ER, the Golgi and throughout the secretory pathway. Hence, the UPR can be thought of as a homeostatic control mechanism to remodel the secretory pathway according to the cell's need.

Ire1p

Ire1p (inositol requiring 1) was originally identified on the basis of its role in inositol prototrophy in *S. cerevisiae* [106]. However, soon thereafter, two other groups linked Ire1p to the UPR, since it was required for the activation of the *KAR2* (BiP) gene upon ER-stress [21,96]. The *S. cerevisiae IRE1* gene encodes a type I transmembrane protein consisting of 1115 amino acids that resides in the ER membrane. Genome wide analysis identified *HAC1* mRNA as the only substrate for Ire1p [107]. Ire1p is the most proximal component of the pathway, and consists of three functional domains. A luminal domain, a serine-threonine kinase domain and a RNase domain. The luminal domain is the most amino-terminal domain that resides in the lumen of the ER and functions as the sensor of the folding state of proteins in the ER [21]. Ire1p senses the load of misfolded proteins (at least partly) through binding with uncomplexed BiP [70,115]. Accumulation of unfolded proteins in the ER results in the release of BiP from Ire1p and causes Ire1p to oligomerise and to undergo trans-autophosphorylation via its cytosolic serine-threonine kinase domain [147,178]. This activates in turn the site-specific endoribonuclease (RNase) domain located at the carboxy-terminus of Ire1p. After being activated, the endoribonuclease domain of Ire1p, in conjunction with Rlg1p, processes the *HAC1* mRNA by splicing an unconventional intron from the messenger [152].

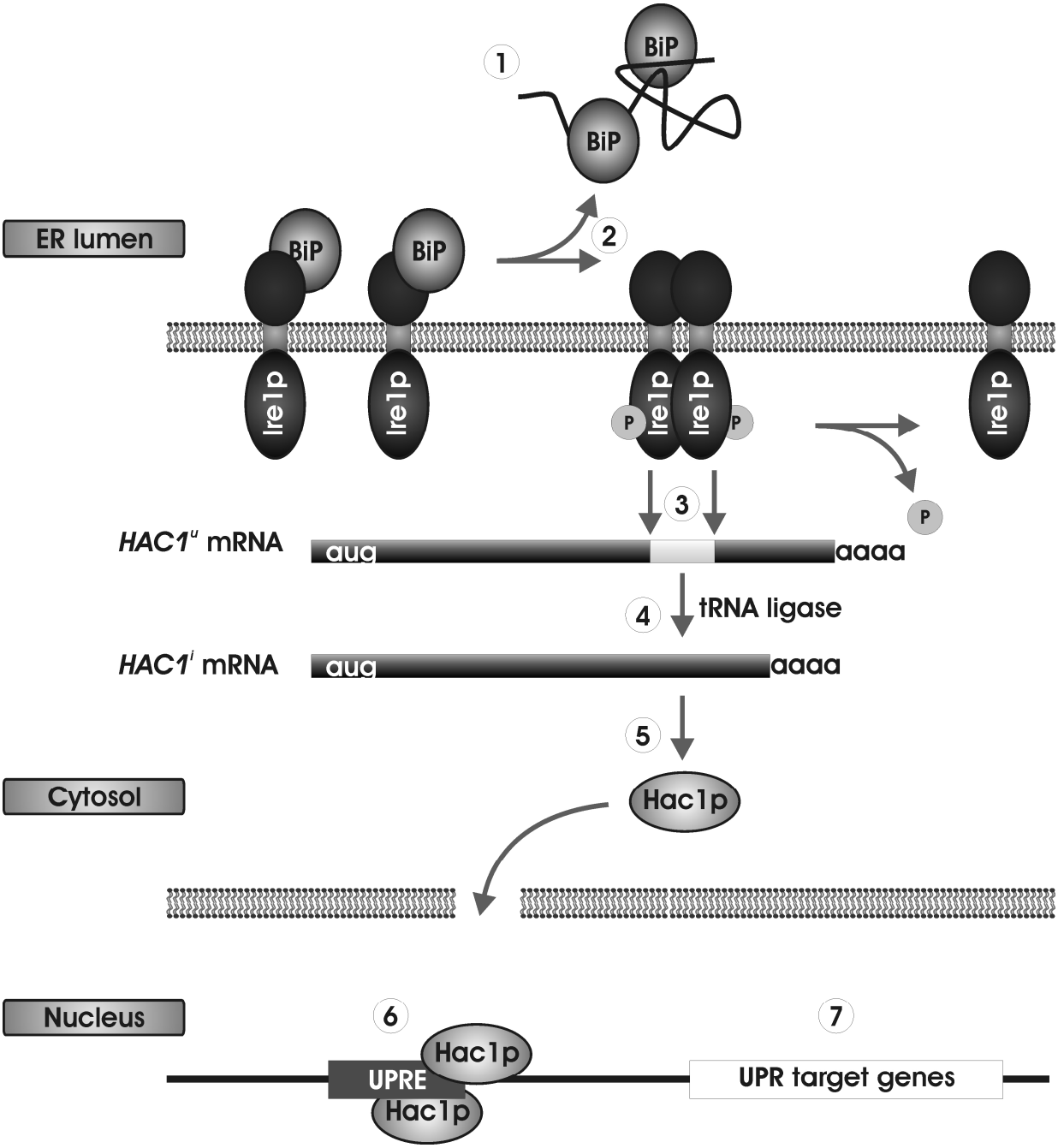


Figure 6. Schematic overview of the UPR in yeast. 1. In a no-stress situation the proximal sensor kinase Ire1p is held in a monomeric inactive state by the chaperone BiP. 2. Upon accumulation of unfolded proteins in the ER BIP dissociates from Ire1p and associates with the unfolded proteins. The released Ire1p undergoes dimerization and autophosphorylation and a site-specific endonuclease (RNase) domain is activated. 3. The RNase cleaves an unconventional intron from the *HAC1* mRNA. 4. The *HAC1* mRNA exon ends are joined together by tRNA ligase. 5. The induced *HAC1* mRNA is efficiently translated to Hac1p. 6. Hac1p translocates to the nucleus and binds to UPR-elements present in the promoters of its target genes. 7. Transcription of genes encoding foldases and chaperones is induced, leading to an increased folding capacity of the ER. 8. When the conditions in the ER are normalized the UPR is switched off by dephosphorylation and depolymerisation of Ire1p conducted by the kinase Ptc2p.

However, dissociation of BiP from Ire1p is not sufficient for activation of the UPR. Deletion of the BiP binding site rendered Ire1p unaltered in ER stress inducibility. It was inactive in unstressed cells but still able to activate the UPR pathway upon ER stress [71,113]. Structural analyses indicated that Ire1p might interact directly with unfolded proteins. Analysis of the crystal structure of the core ER luminal domain (cLD) revealed the presence of a deep groove in cLD dimers, a groove that resembles the major histocompatibility complex (MHC)-like peptide binding groove [23]. The depth of the groove would allow for interactions with unfolded regions of proteins but precludes interaction with compactly folded proteins. Furthermore, mutations in residues lining the bottom of the groove, or at the interface of the dimers abolished Ire1p activation. These findings have led to the hypothesis that Ire1p directly binds unfolded proteins via its cLD. Binding of Ire1p to unfolded regions may promote the dimerisation which is necessary to form the groove. The changes in the luminal domain of Ire1p in turn position Ire1p kinase domains in the cytoplasm optimally for autophosphorylation to initiate the UPR. In this model unfolded proteins act as positive regulators of Ire1p, whereas BiP would act as a negative regulator and would be more an adjuster for sensitivity to stress [71].

The exact localisation of the C-terminal domain of Ire1p has been somewhat unclear. In order to sense the folding state of the ER and transmit the signal to the *HAC1* mRNA, the C-terminal domain of Ire1p must be localised in the vicinity of its substrate (*HAC1* mRNA), and tRNA ligase. Also, Ada5p, which is part of the Gcn5 transcriptional coactivator complex of yeast, which was shown to interact with Ire1p and be required for *HAC1* splicing [180] should be accessible for Ire1p.

The nuclear localisation of Ada5p and that of tRNA ligase [16] together with the assumption that only newly synthesised *HAC1* mRNA would be substrate for Ire1p mediated splicing, led to the hypothesis that the Ire1p catches *HAC1* mRNA during export from the nucleus to the cytosol [151]. Ire1p would thus be localised to the ER and inner nuclear membrane which are contiguous with each other. Its C-terminal half with the endoribonuclease domain would thus face the cytoplasm or the nucleus. The recent finding of a nuclear localisation sequence in the linker region of Ire1p also indicated the nuclear localisation of the C-terminal domain [42]. However, earlier it was also reported that splicing of *HAC1* mRNA precursors that have accumulated on stalled polyribosomes could occur in the cytoplasm [137]. It was therefore suggested that the import of Ire1p into the nucleus is not required for the processing of the preexisting pool of stalled polyribosome-associated *HAC1* mRNA immediately upon induction of ER stress. Only if the UPR continues, nuclear import of Ire1p is essential for the splicing of newly synthesised *HAC1* mRNA precursor in the nucleus [42].

Hac1p

HAC1 mRNA encodes the basic leucine zipper (bZIP) transcription factor that ultimately upregulates the transcription of UPR targets genes. *HAC1* (an acronym for: homologous to *ATF*/*CREB1*) was originally cloned as a multicopy suppressor of the *Schizosaccharomyces pombe* cdc10-129 mutant [110]. Later it was cloned again using a yeast one-hybrid screening, and its role in the UPR was discovered [95]. Under non-stressed conditions, the *HAC1* gene is constitutively transcribed, but the *HAC1*^{u} mRNA ('u for uninduced') is poorly translated due to an unconventional intron near the 3'-end of the transcript [66,137] (Fig. 7A). First after splicing of the inhibitory intron by the action of Ire1p and Rlg1p is the new *HAC1*^{i} mRNA ('i' for induced) efficiently translated to produce the active transcription factor Hac1p [14,66]. Hac1p then translocates to the nucleus where it activates or increases transcription of UPR target genes by binding to UPR specific upstream activation sequences, the unfolded protein response element (UPRE) [22]. The splicing of *HAC1* mRNA is therefore regarded as the key regulatory step in the UPR pathway of *S. cerevisiae*.

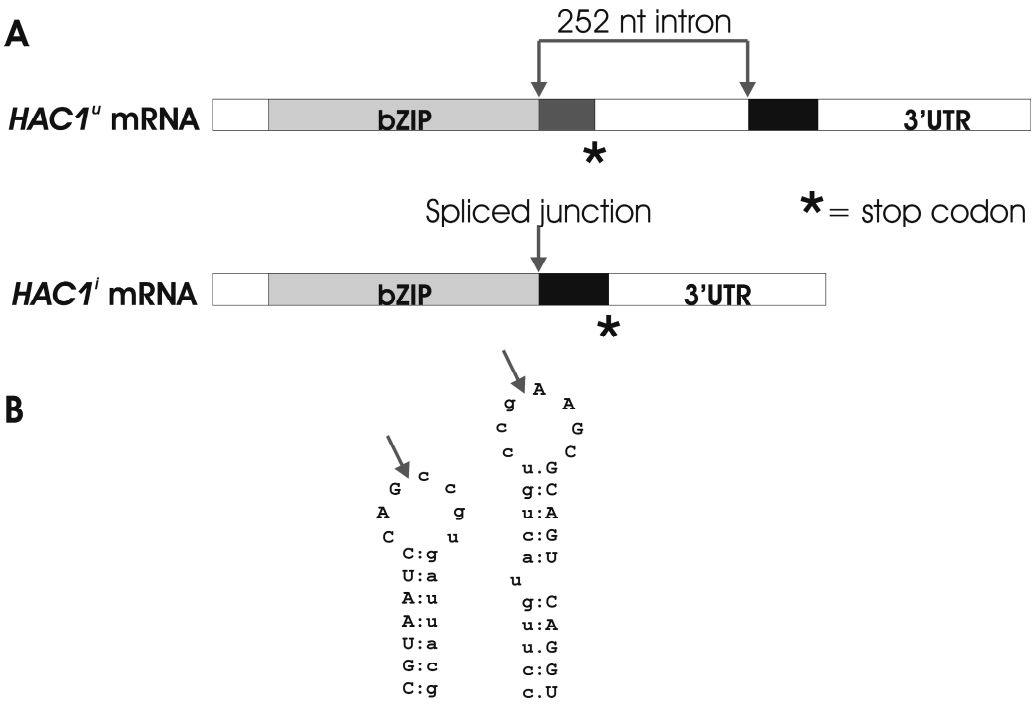


Figure 7. Unconventional intron splicing of the yeast *HAC1* mRNA. **(A)** The 1.4 kb *HAC1u* mRNA contains an ORF encoding a 230 amino acid protein. Splicing of a 252 nt unconventional intron from the *HAC1u* mRNA, changes the C-terminal tail of the encoded protein. The last 10 amino acids of the encoded protein are replaced with an activation domain of 18 amino acids. The dark grey box represents the part encoding the C-terminal 10 amino acids in *HAC1u* mRNA. The black box represents the part encoding the 18 amino acid activation domain. **(B)** The stem-loop structure of the intron borders of *HAC1* mRNA. The Ire1p cleavage sites are indicated by arrows [43].

Spliceosome-Independent Splicing of *HAC1^u* mRNA

The spliceosome-independent splicing of the *HAC1^u* mRNA, mediated by Ire1p and Rlg1p resembles the process of tRNA splicing. Genome-wide analysis identified *HAC1^u* mRNA as the only substrate for Ire1p, suggesting that this unconventional mRNA processing has evolved solely for the signal transduction between ER and nucleus [107]. The unconventional intron of *HAC1^u* mRNA is 252 nt long [22], and the intron borders and splice-sites are organised in two stable stem-loop structures (Fig. 7B). These small stem-loop structures are the substrates for Ire1p and the cleavage at the two splice-sites of *HAC1^u* mRNA occurs independently and in a random order. It requires a well defined 7-nucleotide loop that must be closed by a stem. After excision of the intron by Ire1p the two exons remain associated by base pairing and are ligated by the tRNA ligase Rlg1p [43]. The symmetrical structure of the splice-sites and the fact that Ire1p oligomerises upon activation of the UPR may reflect a situation in which an Ire1p homodimer binds to *HAC1^u* mRNA so that one monomer interacts with the 5'splice-site while the other monomer interacts with the 3'splice-site. Both the *HAC1^i* and *HAC1^u* transcript encode proteins, named Hac1p^i and Hac1p^u respectively. These two proteins have identical N-terminal domains, but differ in both the length and the amino acid sequence of the C-terminal domain (Fig. 7A). The splicing event by Ire1p leads to a frameshift that replaces the last 10 codons of the *HAC1^u* ORF with an exon encoding 18 amino acids [22]. This 18 amino acid tail was found to function as a potent activation domain [98]. The splicing event does not affect the 220 amino acid N-terminal part of the protein that contains both the DNA-binding, and the dimerisation domain. The splicing event therefore functions as a mechanism to join the *HAC1* DNA-binding domain to its activation domain, enabling rapid posttranscriptional generation of a strong transcriptional activator when needed.

Translational Attenuation

Although *HAC1^u* mRNA also encodes a putative bZIP transcription factor, only Hac1p^i accumulates at detectable levels in the cell. It was shown that Hac1p^i and Hac1p^u are equally unstable, and that as a result the lack of Hac1p^u is not due to degradation [66]. Expressing both proteins in engineered yeast strains showed a similar half-life of 2 minutes for both proteins, and indicated a translational control mechanism as the cause for the absence of Hac1p^u [14]. Translation of the *HAC1^u* messenger is initiated, but the ribosomes stall on the unspliced transcript. This translational attenuation is mediated by the unconventional intron. Long-range base pairing of the unconventional intron with the 5'UTR of the *HAC1* transcript blocks the mRNA translation [137]. Part of the 252 nt intron of the yeast *HAC1^u* mRNA forms a stable double-stranded structure by base pairing with the 5'UTR, thereby preventing the ribosomes from reading through (Fig. 8). This long-range base pairing encompasses a stretch of 19 bases with 16 pairings of which 11 are GC pairs.

In the current model [137], the ribosomes start immediate translation of *HAC1*^{u} mRNA as it is exported from the nucleus to the cytoplasm. When the 3'-end of *HAC1* mRNA leaves the nucleus the long-range base pairing comes about, stalling and trapping the ribosomes on the mRNA. Splicing removes the intron and with it one half of the inverted repeat, undoing the long-range base pairing and as a consequence the translational block.

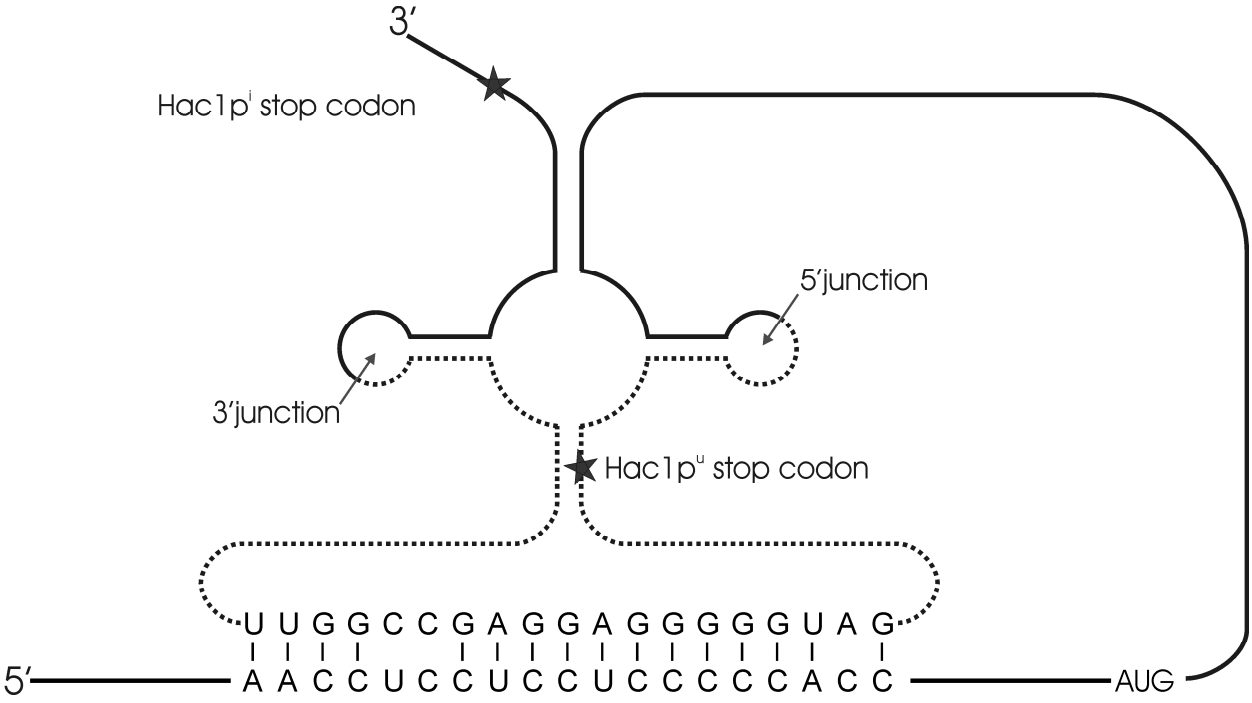


Figure 8. Schematic representation of the secondary structure and long range base pairing in the yeast *HAC1u* mRNA, causing translational attenuation. The *HAC1* 5'UTR and intron contain complementary sequences. Exonic sequences are shown as solid lines and intronic sequences as dashed lines. The intron borders are depicted by the small stem-loop structures, and the splice-sites indicated by the arrow. Modified after Rüegsegger *et al.* [137].

The Unfolded Protein Response Element (UPRE)

After splicing of the unconventional intron from the *HAC1*^{u} mRNA translation resumes, and the resulting Hac1p^{i} protein translocates to the nucleus where it binds to the UPRE. UPREs are *cis*-acting regulatory elements functioning as upstream activator sequences (UAS) or UPR-regulated genes. The UPRE was originally identified as a 22 bp *cis*-acting element present in the yeast KAR2 gene (BIP), and was found necessary and sufficient for the transcriptional activation in response to ER-stress [72,99]. Later on, mutational analysis revealed the core sequence of the yeast KAR2 UPRE, a semi palindromic seven nucleotide sequence with a spacer nucleotide between both half-sites (**CAGCGTG**) [95]. This one nucleotide spacer is important for binding of the UPRE by Hac1p. The lack of spacing reduced the induction of a reporter construct to 13% of the normal level, whereas increased spacing abolished the induction almost completely. The core sequence of the UPRE is well conserved and has also been found in the promoters of other UPR regulated genes. It was further shown that Hac1p shows a preference for a specific spacer nucleotide in the order **C > G > A > T** [97].

Additonal Regulation in the Yeast UPR

In addition to the simplified model of the UPR schematised in Figure 6, several other mechanisms have been discovered which turn the simple on-off switch concept of the yeast UPR into a more complex response.

Promoter analysis of the *HAC1* gene itself revealed the presence of the UPRE-like sequence **CACCTTG**. It was shown that Hac1p activates its own transcription by binding to this *cis*-acting element present in its own promoter, which was necessary and sufficient for the induction of the *HAC1* gene by ER stress [112]. Cells that lack this form of autoregulation cannot maintain high levels of *HAC1* mRNA, and thereby not survive very well under prolonged ER-stress. It is therefore likely that this autoregulation is required for sustained activation of the UPR. Furthermore, an Ire1p independent pathway was reported for *S. cerevisiae*, which adds additional control to the UPR pathway [79]. Transcription of the *HAC1* gene could be induced in an Ire1p-independent manner, and is as a result different from the autoregulation described above. The Ire1p independent *HAC1* induction requires a bipartite signal (Fig. 9), consisting of one input by unfolded proteins in the ER (UP), and the other input provided by inositol starvation or temperature shift (ino/temp), conditions that both affect the ER membrane properties. ER membranes distressed by either inositol deprivation or elevated temperature might control the activity of a membrane bound component of a signal transduction machine that also senses protein folding conditions in the ER lumen. The Ire1p independent transcriptional induction of *HAC1* combined with Ire1p mediated splicing results in elevated Hac1p levels, and is referred to as 'Super-UPR' (S-UPR).

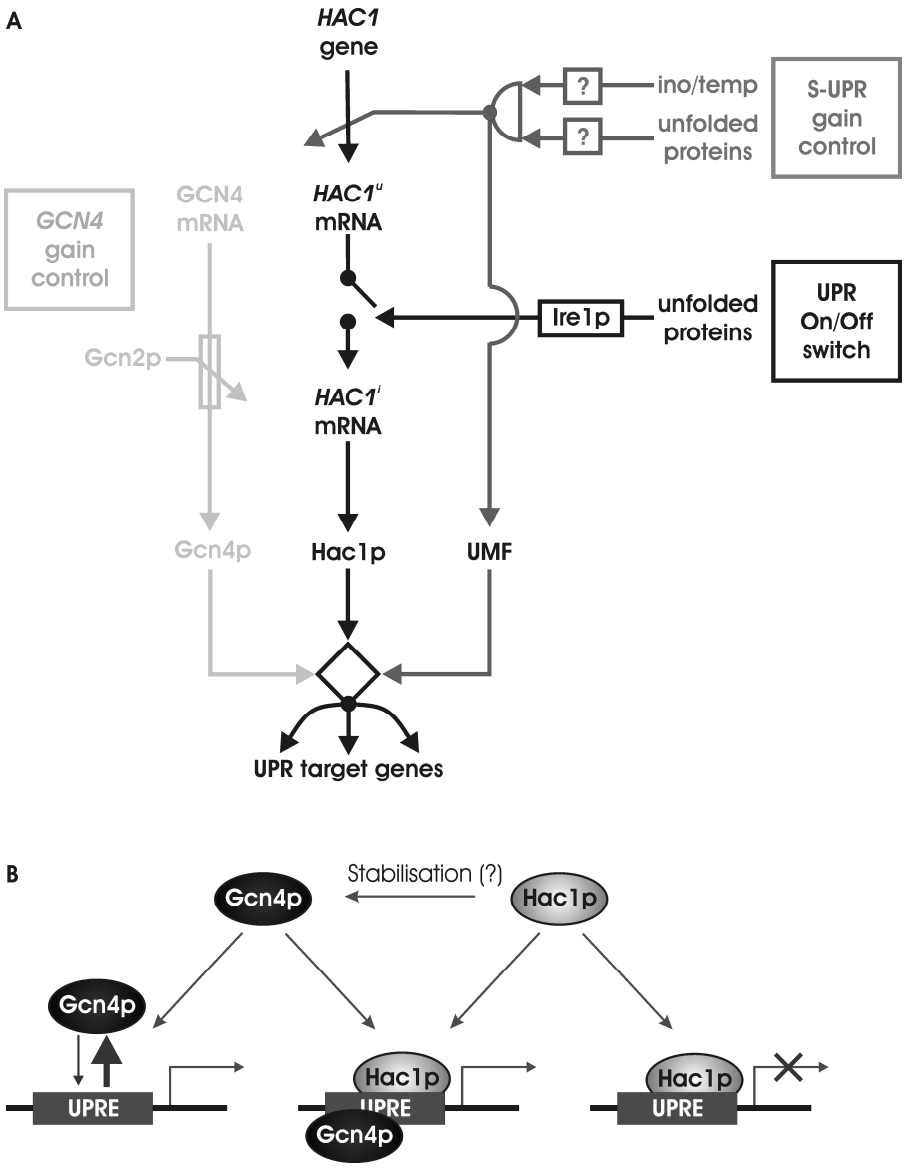


Figure 9. Model of the expanded circuitry of the yeast UPR.
(A) The classical UPR is represented in black, the S-UPR in dark grey and the GCN4 branch in light grey. The three branches of the circuitry are integrated at target promoters. Transcriptional regulation of HAC1 mRNA levels, providing one level of gain control, is depicted as a rheostat under supervision of a logical AND gate informed by multiple inputs from the ER. Splicing of HAC1 mRNA by Ire1p, is depicted by a binary on/off switch. Regulation of Gcn4p levels by Gcn2p under changing cellular conditions adds an additional layer of gain control. Together, activity levels of Hac1p, Gcn4p, and the proposed UPR modulating factor [79] collaborate to determine the magnitude of the transcriptional output signal. **(B)** Mechanism of Gcn4p/Hac1p action at target promoters. In the absence of Hac1p, Gcn4p is present in the cell as a consequence of baseline activity of Gcn2p. At normal concentrations, Gcn4p is unable to bind or activate a target UPRE, but it may bind when Gcn4p levels are elevated. Upon induction of the UPR, Ire1p is activated and Hac1 is synthesised. Hac1p can bind, but not activate, target UPREs. Binding of target DNA by a Gcn4p/Hac1p heterodimer, results in a transcriptionally active complex. Gcn4p levels are upregulated under UPR induction, perhaps as a consequence of stabilisation by interaction with Hac1p. Modified after Patil *et al.* [124]

Although the molecular details of the S-UPR are not known, it was proposed that it is most likely the result of the production or activation of an additional transcriptional activator, termed the UPR modulatory factor (UMF) [79]. The output of the S-UPR in yeast was found not to be simply proportional to the Hac1p concentration. Promoters of target genes rather displayed a differential responsiveness to the Hac1p concentration and UMF activity. Some target genes were already fully induced during normal UPR, whereas others were further induced by S-UPR. Therefore, the promoters of target genes possibly have differential affinity for Hac1p and/or UMF.

Besides the UMF, additional UPREs and the transcription factor Gcn4p play a role in the *S. cerevisiae* UPR. Promoter analysis of the 381 genes identified as UPR target genes [160], failed to identify a *KAR2* like UPRE in most of them. However, computational analysis of promoters of UPR target genes for overrepresented short upstream activating sequences (UASs) led to the identification of two novel UPREs (UPRE-2: TACGTG and UPRE-3: AGGACAAC) [124]. Similar to the earlier identified *KAR2* UPRE (now called the UPRE-1), these were both shown to be necessary and sufficient to confer UPR inducibility on a target gene. A genetic screen for activators of these newly identified motifs revealed that Gcn4p plays a role in the UPR. Gcn4p is a bZIP transcriptional activator and it has been characterised as a component of the cellular response to amino acid starvation and other stresses [102].

It was shown that Gcn4p and Hac1p act together at target gene promoters. Over-expression of Hac1p did not activate transcription of target genes in the absence of Gcn4p, whereas Gcn4p in the absence of Hac1p was also insufficient to induce transcription [124]. Gcn4p and Hac1p are both bZIP type transcription factors and therefore possibly act as hetero-dimer complexes. The role of Gcn4p in the UPR circuitry of *S. cerevisiae* is schematised in Figure 9.

The UPR in Higher Eukaryotes

The UPR in higher eukaryotes shares many similarities with the yeast UPR, but the mechanism seems more complex than in yeast. From the higher eukaryotes, especially the mammalian UPR has been studied intensively, as many human diseases are the results of altered protein quality control in the ER [116]. The major difference with the yeast UPR is the presence of three ER-localised transmembrane signal transducers that are activated to initiate the UPR. These three signal transducers in mammalian cells are: the yeast Ire1p homologue IRE1 [186], a protein kinase-like ER kinase (PERK) [52], and the activating transcription factor 6 (ATF6) [54,185]. IRE1, PERK and ATF6 are constitutively expressed in metazoan cells, all three sense unfolded proteins via an interaction with BiP. BiP is therefore regarded as the master regulator of the UPR in mammalian cells. In unstressed cells the luminal domains of PERK, IRE1 and ATF6 are associated with BiP, and kept in an inactive state. In response to ER stress, unfolded proteins accumulate and bind BiP, thereby

sequestering BiP and promoting the release of BiP from the luminal domains of the three sensors [8]. A schematic overview of the three branches in the UPR of mammalian cells is given in Figure 10.

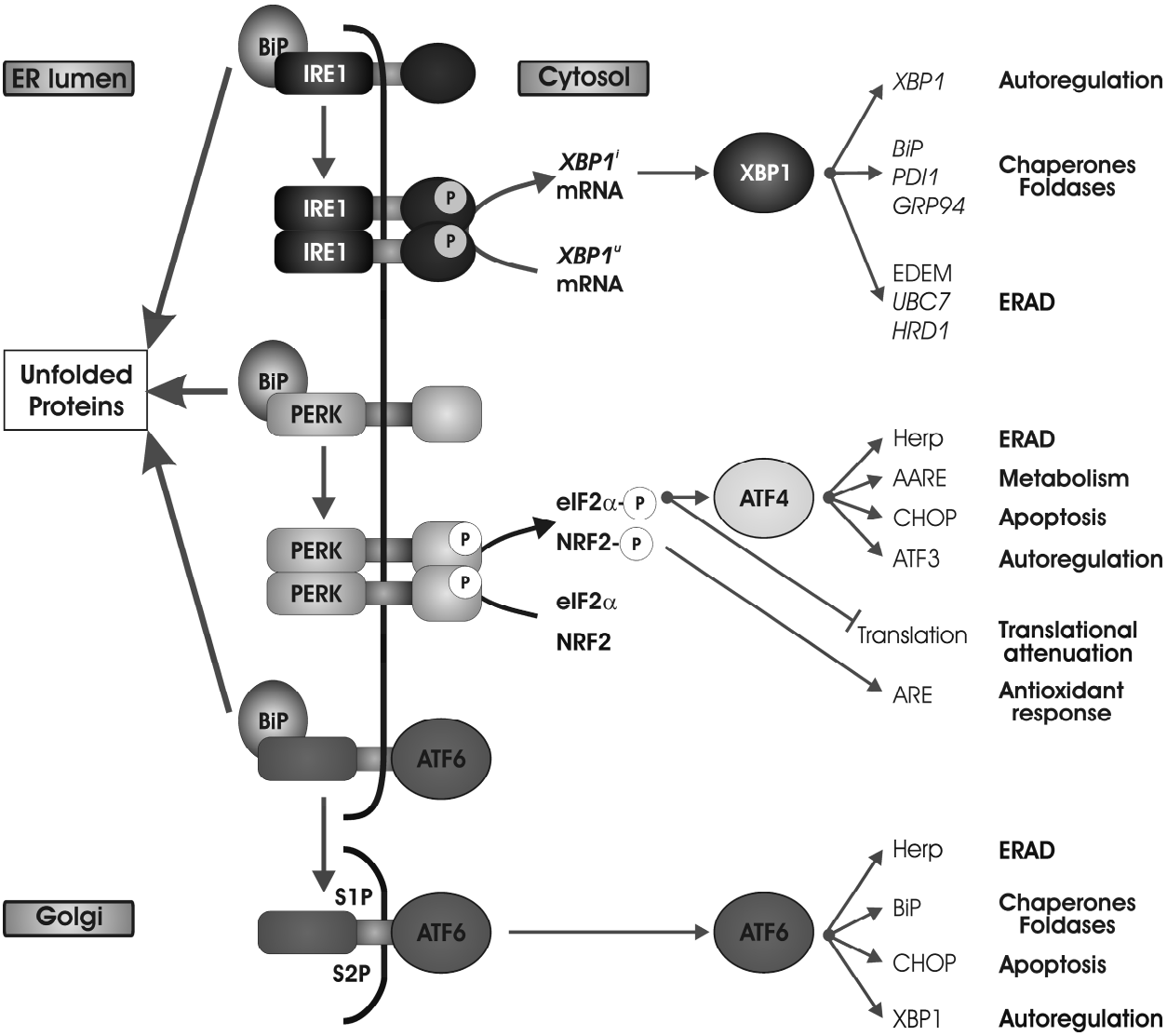


Figure 10. Model for the mammalian unfolded protein response. The mammalian ER contains three types of transmembrane proteins (IRE1, PERK, and ATF6), which sense the presence of unfolded proteins in the ER and transmit signals across the membrane. ER stress induced activation of the IRE1-XBP1 pathway (top), PERK-eIF2a phosphorylation-ATF4 pathway (middle), or ATF6 pathway (bottom) culminates in induced transcription of respective target genes. Abbreviations: AARE, amino acid response element; ARE, antioxidant response element; CHOP, CCAAT/enhance binding protein (C/EBP) homologous protein; NRF2, nuclear factor erythroid 2 (NF-E2) related factor; UBC7, ubiquitin conjugating enzyme 7; XBP1, X-box binding protein; HERP, homocysteine-induced ER protein.

PERK

Upon ER stress the most immediate response in mammalian cells is the attenuation of mRNA translation, which lowers the influx of newly synthesised proteins into the ER [65]. This translational attenuation is mediated by PERK. Like IRE1, PERK is an ER-associated transmembrane serine/threonine protein kinase, and the luminal domains of IRE1 and PERK regulate the protein kinase of these proteins by functioning as ER stress regulated oligomerisation domains. The function of the luminal domains of IRE1 and PERK is also conserved across species as was shown by a chimeric Ire1p. The luminal domain of the yeast Ire1p could be functionally replaced with the luminal domain of either IRE1 or PERK [84].

ER stress leads to PERK dimerisation and *trans*-autophosphorylation, which results in activation of PERK [52]. The activated PERK phosphorylates the alpha subunit of the eukaryotic translation initiation factor 2 (eIF2α) on serine residue 51, which leads to its inactivation and reduces translational initiation.

Using a mathematical model it was shown that adding a translation attenuation mechanism to the UPR signalling network reduces the amount of free unfolded proteins, and results in fewer excess chaperones during consecutive stresses [161]. This is the result of the translational attenuation response being fast compared to chaperone upregulation.

In addition to global repression of translation, activation of PERK also induces transcription of a number of UPR dependent genes [50,53,143]. Phosphorylation of eIF2α inhibits general translation, with the exception of several mRNAs for which it is required for their selective translation. For instance, phosphorylation of eIF2α induces translation of the mRNA encoding the activating transcription factor 4 (ATF4) [51]. ATF4 is required for expression of genes involved in amino acid biosynthesis, anti oxidative stress responses, and apoptosis.

Besides eIF2α, the transcription factor nuclear respiratory factor 2 (NRF2) is a substrate for PERK kinase activity [24]. NRF2 is involved in the oxidative stress response, and in unstressed cells, NRF2 is kept latent in cytoplasmic complexes via association with the cytoskeleton anchor, Keap1 (kelch-like ECH-associated protein1). PERK dependent phosphorylation of NRF2 is necessary and sufficient to trigger dissociation of the NRF2/Keap1 complexes. This allows nuclear import of NRF2 where it promotes expression of its downstream target genes [61].

IRE1 and XBP1

In mammalian cells two homologues of the yeast Ire1p are present, IRE1α and IRE1β [159,175]. IRE1α is expressed in most cells and tissues, whereas IRE1β is expressed predominantly only in intestinal epithelial cells. The function of the mammalian IRE1 is analogous to the

yeast Ire1p [108]. Upon accumulation of unfolded proteins in the ER, its activated endoribonuclease domain mediates the stress-induced splicing of its substrate, *XBP1* mRNA (X-box binding protein) [12,148]. Like its yeast homologue Hac1p, XBP1 is a bZIP transcription factor and its mRNA contains a 26 nucleotide unconventional intron. The predicted secondary structure for this intron shows similarity with the predicted structure for the *hacA*^{u} mRNA. The exon-intron junctions are encompassed in two loops of seven unpaired nucleotides (see Fig. 3, Chapter 2). The removal of the 26 nucleotide unconventional intron from *XBP1*^{u} mRNA creates a frameshift, and as a result, the spliced *XBP1*^{i} mRNA encodes a larger form of XBP1 that contains a novel transcriptional activation domain at its C-terminus [81,186]. XBP1 regulates a subset of ER resident chaperones including some DnaJ-like accessory proteins, oxidoreductases, and PDI-P5 [80]. It also has a modest effect in regulating BiP, ATF6 and itself.

XBP1 was isolated (together with ATF6) using the mammalian ER stress response element (ERSE-I) in a yeast one-hybrid screen [185]. The ERSE-I was earlier identified by transcriptional analysis of UPR target genes such as BiP, GRP94 (glucose regulated protein 94) and calreticulin. ERSE-I has a tripartite structure with the consensus sequence CCAAT-N_9-CCACG [136]. Both XBP1 and ATF6 bind the element in a complex with the general transcription factor NF-Y, which binds to the CCAAT part of the ERSE-I [188]. The spacing of the two motifs was shown to be critical for binding. ATF6 binds CCACG only when CCAAT is exactly 9 base pairs upstream and bound by NF-Y [187]. In addition to ERSE-I, two additional *cis*-acting elements have been identified, the ERSE-II and the UPRE.

ERSE-II, which resembles the ERSE-I, was identified in the promoter region of the *HERP* gene (homocysteine-induced ER protein). *HERP* encodes an ER-membrane protein, is highly upregulated during ER stress [73], and enhances ERAD [69]. The consensus sequence of the ERSE-II is ATTGG-N-CCACG [74]. Like ERSE-I it also contains two motifs, ATTGG (complementary to CCAAT) and CCACG, although the interval between them is different. ATF6 binds to the ERSE-II in a NF-Y dependent manner, as is the case for ERSE-I. XBP1 on the other hand could bind the ERSE-II in a NF-Y independent manner [183].

The mammalian UPRE was identified as a binding site for ATF6 and has a consensus sequence of TGACGTG(G/A) [176]. Later, however, it was shown that ATF6 prefers NF-Y dependent binding to the half-site (CCACG) in the ERSE-I to binding to the UPRE by itself, whereas XBP1 binds both to the ERSE-I and UPRE [186].

ATF6

The third sensor of unfolded proteins in the ER of mammalian cells is the bZIP-containing transcription factor ATF6, which is initially synthesised as an ER membrane spanning protein [54,185]. In unstressed cells, ATF6 is associated with BiP which keeps it localised to the ER membrane [8,15]. ER stress induces dissociation of BiP and permits the trafficking of ATF6

to the Golgi where it is cleaved by two proteases [184,187]. Site-1 protease (S1P, a serine protease) cleaves ATF6 in its luminal domain, and the N-terminal membrane-anchored half is cleaved by site-2 protease (S2P, a metalloprotease) within the phospholipid bilayer. These proteolytic reactions release the cytoplasmic domain bZIP domain of ATF6, which than can translocate to the nucleus and bind ERSEs to activate its target genes.

The UPR *A. niger*

Despite the exploitation of the secretory pathway of fungi for the commercial production of enzymes, the understanding of the pathway and the cellular mechanisms behind it has relatively long remained a black box. Mostly, knowledge from other eukaryotic organisms was extrapolated to fungi [125]. The first information regarding the UPR in filamentous fungi came from the isolation of the genes encoding BipA, PdiA and TigA, and the finding that the expression of those chaperones and foldases was upregulated in response to secretion stress [58,62,104,130]. Later, the functional homologue of the yeast Hac1p was cloned from *Trichoderma reesei, A.nidulans* [140], and from *A. niger* [101], and soon after the genes encoding the Ire1 and Ptc2 were characterised in *T. reesei* [165] enabling the study of the fungal UPR in more detail. Later, all key elements constituting the signal transduction pathway of the yeast UPR (Fig. 6) have been identified in filamentous fungi, suggesting that the pathway resembles the *S. cerevisiae* UPR. Indeed, the *A. niger* pathway shares common features with the yeast UPR, and the central activation mechanism of the pathway triggered by unconventional intron splicing is conserved among yeast, filamentous fungi and mammalian cells. There are, however, some unique features in the UPR of *A. niger* which are the subject of study in this thesis and will be discussed later.

The fact that the UPR of filamentous fungi is an important target to investigate with respect to the improvement of (heterologous) protein production was illustrated in a study by Valkonen *et al.* [166] soon after the isolation of the fungal *hac1/hacA* genes. They over-expressed the spliced form of *hacA*, thereby constitutively activating the pathway downstream of HacA, and showed that the constitutive activation of the *A. niger* UPR enhanced the production of *Trametes versicolor* laccase and bovine chymosin in the fungus several fold.

Repression Under Secretions Stress (RESS)

In mammalian cells, attenuation of general translation initiation mediated by PERK diminishes the protein load in the ER upon folding stress. Although filamentous fungi do not appear to contain a PERK homologue, a different mechanism accompanies the UPR which also diminishes the pool of newly synthesised proteins. Under secretion stress conditions, *T. reesei* and *A. niger* specifically down regulate transcription of secretory

proteins by a mechanism termed RESS (Repression under Secretion Stress) [1,119]. RESS minimises hereby the input of proteins into the ER and this may be a different way to decrease the load of the ER, as compared to the PERK-dependent pathway in mammalian cells. Although the exact mechanism by which RESS operates is unknown, it has been shown to take place at the transcriptional level, and to be mediated through the promoter of the gene involved. Truncation of the *T. reesei cbh1* (cellobiohydrolase 1) promoter abolished its transcriptional down-regulation by RESS [119], and for the *A. niger glaA* gene (glucoamylase) the promoter region between 1 and 2kb upstream was shown to be important for RESS [1]. Even though RESS commonly coincides with the UPR, it is not necessarily activated to a similar degree; a mild UPR might coincide with a strong RESS. Al-Sheikh *et al.* [1] investigated the possible involvement of HacA in RESS. They observed RESS in an antisense *pdiA* strain of *A. niger* where there was no apparent induction of the UPR, indicating that RESS might be mediated separately from the UPR. The possibility that this involves an ER-to-nucleus signal transduction pathway independently from the IreA branch still needs to be examined.

The UPR and Genomics

The recently published genome sequences of several *Aspergilli* [37,87,126] have made genomics approaches feasible to investigate the responses due to protein secretion in those filamentous fungi.

Using transcriptomics, Sims *et al.* [154] studied the effects of a heterologous protein in *A. nidulans* by comparing a chymosin producing strain with its parental wild type strain. Of the genes that were identified as significantly upregulated in the chymosin producing strain, 22 were involved in UPR and/or protein folding (including chaperones and foldases like *bipA, clxA pdiA, prpA* and *tigA,* several heatshock proteins, and *hacA* itself). Furthermore, genes with functions in vacuolar protein sorting, protein degradation (ERAD), vesicle transport, lipid and inositol metabolism, cellular export and secretion, and cell wall biogenesis were found to be specifically upregulated by chymosin expression, indicating that the UPR of *A. nidulans*, like the yeast UPR [160] is a broad response affecting numerous cellular processes.

Arvas *et al.* [4] compared the transcriptional response to secretions stress in *T. reesei* with the transcriptomics data on various stress responses in *S. cerevisiae*. The genes upregulated in response to secretion stress included a large number of secretion related genes in both organisms, indicating the similarities between the stress responses in both organisms. In addition, however, they found that the *T. reesei cpc1* gene was upregulated in response to secretion stress, in contrast to its homologue in yeast. Transcription of *cpc1* was induced in DTT treated cultures as well as in *T. reesei* cultures expressing human t-PA (tissue plasminogen activator) or over-expressing *ire1* indicating its regulation by UPR.

cpc1 is the homologue of the yeast GCN4 that encodes a bZIP transcriptional activator involved in the cellular response to amino acid starvation and other stresses [102], and plays a role in the yeast UPR (Fig. 9), but has not been shown to be induced by secretion stress.

The first genome wide expression analysis of secretion stress related responses in *A. niger* was conducted by Guillemette *et al.* [48]. In their study, ER stress was triggered either by DTT or tunicamycin treatment, or by the expression of a heterologous protein (t-PA). It was found that each condition affected its own set of genes and that only ten genes were upregulated in all three conditions, which were mostly ER localised chaperones and foldases. Using polysomal fractionation they also showed that *hacA*^{u} was present in the non-polysomal fractions, suggesting that translation of *hacA*^{u} mRNA might be blocked. In this thesis, evidence is presented for such a translational block, which is present only in the *hacA*^{u} transcript. Their results also pointed out the possible existence of a feed-back mechanism at the translational level. mRNAs encoding several secreted proteins, including glucoamylase, were redistributed from polysomes to monosomes after DTT treatment, indicative of translational repression. Thus, in addition to RESS, *A. niger* might possess another feed-back mechanism aiming at reducing the amount of newly synthesised proteins translocated into the ER lumen upon secretion stress.

In conclusion, the response of *A. niger* to the accumulation of unfolded proteins in the lumen of the ER might involve four mechanisms to suppress the stress that poses on the fungus; 1) boosting the capacity of the secretory pathway by increasing expression of genes involved in protein folding and transport. 2) increasing the ER associated degradation (ERAD) to eliminate terminally unfolded proteins from the secretory pathway. 3) decreasing transcription of secretory proteins to reduce the load of newly synthesised proteins into the ER (RESS). 4) translational repression of secretory proteins to lower the load of newly synthesised proteins translocated into the ER in addition to RESS.

Aim and Outline of this Thesis

The aim of this thesis is to investigate the unfolded protein response in *A. niger*, and to characterise components of this signalling pathway. Understanding of the UPR is of importance for heterologous protein production in *A. niger*. The fungus is capable of secreting large amounts of a wide variety of enzymes that in nature are needed to release nutrients from complex biopolymers. This characteristic is exploited by the industry for the production enzymes, both in solid state and submerged fermentations. However, whereas the production of homologous proteins can reach up till 40 g/l, the production of heterologous proteins is often several orders of magnitude lower. Expression of heterologous proteins often results secretion stress and activation of the UPR. Understanding of the UPR and other processes in the secretory pathway might ultimately lead to the identification of bottlenecks enabling strategies to optimise of heterologous protein production.

Chapter 2 describes the isolation of *hacA*, the gene encoding the transcription activator of the *A. niger* UPR. It is shown that different forms of ER stress result in the same response; the effect of expression of t-PA or treatment of the mycelia with either tunicamycin or DTT is manifested by the induction of foldases and chaperones and the processing of *hacA* mRNA. It is also shown that a 20 nt unconventional intron is spliced from the *hacA* messenger upon ER stress, and that the 5'UTR looses 230 nt upon secretion stress. Finally it is shown that *hacA* up-regulates its own transcription and that purified HacA protein can bind to promoter fragments of genes encoding foldases and chaperones. **Chapter 3** describes the identification and characterisation of the UPRE. It is shown that the UPRE is a partly palindromic sequence, separated by a single nucleotide and flanked by a highly conserved 4 nt stretch. Furthermore, an optimal binding site for HacA is isolated using an *in vitro* selection procedure, and it is shown that the *hacA* promoter itself contains 3 UPREs of which one resembles the *in vitro* selected binding site. In C**hapter 4** the regulation of *hacA* is investigated. The promoter of *hacA,* and the 5'UTR of the *hacA* mRNA contain various elements, and influence of those elements on the regulation of *hacA* has been studied. The shortening of the 5'UTR of *hacA* mRNA by 230 nt, as described in Chapter 3, is here shown to be due to ER stress induced transcription of *hacA* from a new start site closer to the ATG. This transcriptional switch is operated by HacA and dependent on UPRE2, and relieves the new transcript from a translational block. In C**hapter 5** the results from this thesis are summarised and discussed.

Chapter 2

The Transcription Factor HacA Mediates the Unfolded Protein Response in *Aspergillus niger*, and Up-Regulates its Own Transcription

Harm J. Mulder, Markku Saloheimo, Merja Penttilä, Susan M. Madrid

Published in Molecular Genetics and Genomics. 2004, 271(2): 130-140

Abstract

The unfolded protein response (UPR) involves a complex signalling pathway in which the transcription factor HacA plays a central role. Here we report the cloning and characterisation of the *hacA* gene and its product from *Aspergillus niger*. Endoplasmic reticulum (ER) stress results in the splicing of an unconventional 20 nt intron from the *A. niger hacA* mRNA, and in addition to the truncation of the 5'-end of the *hacA* mRNA by 230 nt. In this study the UPR was triggered by over-expressing tissue plasminogen activator (t-PA), and by treatment of mycelia with dithiothreitol (DTT) or tunicamycin. Over-expression of the processed form of *hacA* not only led to the upregulation of *bipA*, *cypB* and *pdiA*, mimicking an UPR, but also led to the upregulation of the *hacA* gene itself. *In vitro* binding assays confirmed that the HacA protein binds to the promoters of genes encoding ER-localised chaperones and foldases, and to the promoter of *hacA* gene itself. Finally, a GFP-HacA fusion was shown to localise in the nucleus.

Introduction

The endoplasmic reticulum (ER) serves as a folding compartment for membrane and secretory proteins. In the oxidising environment of the ER, protein folding is assisted by numerous protein chaperones and foldases [39], and a variety of quality control mechanisms ensure that only correctly folded proteins are released [30,190]. A variety of physiological or environmental stress conditions can disturb protein folding and lead to the accumulation of unfolded proteins in the ER. Cells respond to such an accumulation by increasing the transcription of genes that code for ER-resident chaperones and foldases, thereby increasing the folding capacity of the compartment. The signalling pathway involved is part of the unfolded protein response (UPR) [64,123,150,179]. Genome-wide analysis in yeast, using DNA micro arrays, has linked a large number of genes to the UPR, with functions far beyond those ER chaperones [38,160]. By affecting virtually every stage of the secretory pathway, the UPR appears to be a versatile regulatory system that serves to maintain the homeostasis of ER functions under stress conditions. The key step in the activation of the UPR in yeast is the splicing of an unconventional intron from *HAC1* mRNA, which serves to relieve a translational block. [66,137]. *HAC1* mRNA encodes the basic leucine zipper (bZIP)-type transcription factor Hac1p, which directly activates the transcription of UPR genes. Splicing of the yeast *HAC1* mRNA occurs in a non-spliceosomal manner. During ER stress, the intron borders of *HAC1* mRNA are cleaved by Ire1p [43,152], after which the exons are rejoined by the tRNA ligase Rlg1p [151]. Ire1p, a kinase/RNase which resides in the ER

membrane senses the accumulation of unfolded protein via a dynamic interaction with the chaperone BiP [8]. Release of BiP triggers the dimerisation and subsequent trans-phosphorylation of Ire1p, thereby activating the UPR. In animal cells, nearly all the features of the yeast UPR splicing system are conserved, but in comparison with yeast their ER stress pathway appears to be more extensive [148,186]. The transcription factor XBP1, which is involved in the ER stress response of *Caenorhabditis elegans* and mammalian cells, is activated through intron splicing mediated by IRE1, in a manner similar to that seen in yeast. In addition, transcription of the *XBP1* gene is induced by ATF6, an ER membrane associated transcription factor, which is proteolytically released during ER stress [81].

In this study we focus on the unfolded protein response in the filamentous fungus *Aspergillus niger*, an industrial production organism which is capable of secreting large amounts of native proteins like glucoamylase. Yields of heterologous proteins, however, are often low [2]. In order to improve *A. niger* as a host for homologous and heterologous protein production, a detailed knowledge about the functioning of the secretion pathway and its bottlenecks is important. Here we report the isolation and characterisation the *A. niger hacA* gene, the functional homologue of the yeast *HAC1*, which codes for the transcriptional activator of the UPR.

Materials and Methods

Strains, Culture Conditions and Transformations

Escherichia coli strains TOP10 and DH5α (Invitrogen) were used as hosts to propagate plasmids. The cDNA library from *Aspergillus niger* var. *awamori* strain UVK143f was constructed in the vector pYES2 (Invitrogen). *A. niger* N592 (*cspA1, pyrA5*) was used as the recipient strain for transformation, and *A. niger* N402 (*cspA1*), and *A. niger* D15::p*glaA*-t-PA#19, which produces the human tissue plasminogen activator protein (t-PA) [181], were used to investigate the UPR under different stress conditions.

All *A. niger* strains were grown on minimal media (MM) plates containing per litre 6 g $NaNO_3$, 1.5 g KH_2PO_4, 0.5 g $MgSO_4 \cdot 7\ H_2O$, 0.5 g KCl, trace elements as described by Vishniac and Santer [172] and 20 g fructose as a carbon source (pH 6.0). Liquid cultures were grown in MM supplemented with 2 g casamino acids and 5 g yeast extract (CM), and inoculated with 10^6 spores/ml and incubated at 30°C. All growth experiments for subsequent Northern analysis were started in fresh CM media with 1.5 g mycelia of an overnight culture. Except for *A. niger* D15::p*glaA*-t-PA#19, which was transferred to MM with 20 g maltodextrin, for induction of the glucoamylase promoter. ER stress was provoked by adding either 3 mg/ml dithiothreitol (DTT) or 10 µg/ml tunicamycin. All t_0 samples were taken just before the transfer to the experimental medium.

A. niger protoplasts were co-transformed according to the procedure of Kusters–van Someren *et al.* [76] with (1) plasmid pHM50, which expresses the coding region of *hacA* without the 20 nt unconventional intron (*hacA*^{i}) under the control of the constitutive *gpdA* promoter and the *trpC* terminator, and (2) pGW635, which contains the *pyrA* gene of *A. niger* [45]. pHM52Δ4, which expresses a *gfp-hacA*^{i} fusion under control of the *gpdA* promoter, also contains the *pyrA* gene, thus allowing direct transformation of *A. niger* N592.

Gene Isolation

RNA was extracted from *A. niger* var. *awamori* strain UVK143f after growth of mycelium on glucose, starch or beet pulp as a sole carbon source, and from mycelium grown under nitrogen starvation. The RNAs were mixed and used for library construction [174]. Southern hybridisation and screening of the cDNA library by colony hybridisation were performed as described by Sambrook *et al.* [141]. The library colony hybridisation was done overnight in a hybridisation mix without formamide [141] at 58°C. Subsequently the filters were washed in 2xSSC, 0.1% SDS for 2x 5 minutes at room temperature, and with the same solution at 58°C for 30 minutes. PCR based genomic walking was performed as described by Siebert *et al.* [153] to isolate the flanking regions of the *hacA* gene. DNA sequencing reactions were performed with the Bigdye sequencing kit (Applied Biosystems) and analysed with the ABI Prism 3100 Genetic Analyser (Applied Biosystems).

Tabel I: Oligonucleotide primers used in this study

Primer	Sequence	Site
P1	5'-ACCATGGTGGAAGAAGCATTCTCTCCAG	*NcoI*
P2	5'-CCGCTGCAGGATGTTGTGTCA	*PstI*
P3	5'-GACCTGCAGTGTCCGTCGCTG	*PstI*
P4	5'-GGGATCCTAACAGCCAGCTGCAATGCCCTG	*BamHI*
P5	5'-GGGATCCTAGTGGTGGTGGTGGTGGTGACAG CCAGCTGCAATGCCCTG	*BamHI*
P6	5'-GGGCGCCTCCATGATGGAAGAAGCATTCTCTC	*NarI*
P7	5'-GCCATGGTGAGCAAGGGCGAGGAG	*NcoI*
P8	5'-CGGCGCCGGACTTGTACAGCTCGTCCATGCC	*NarI*
P9	5'-CGTCGAGAACGTCAAAGGCGAACCCGTC	

Plasmid Construction

For over-expression of *hacA* without the need for a stress induced splicing event, and for construction of a *gfp-hacA*^{i} fusion, the *hacA* ORF was amplified in two fragments from *A. niger* cDNA by PCR. Joining of the two gene fragments at the *PstI* site resulted in reconstitution of the *hacA* ORF without its unconventional intron. The primers used in this study are listed in Table 1, and the positions of the primers used to amplify *hacA*^{i} are indicated in Figure 1.

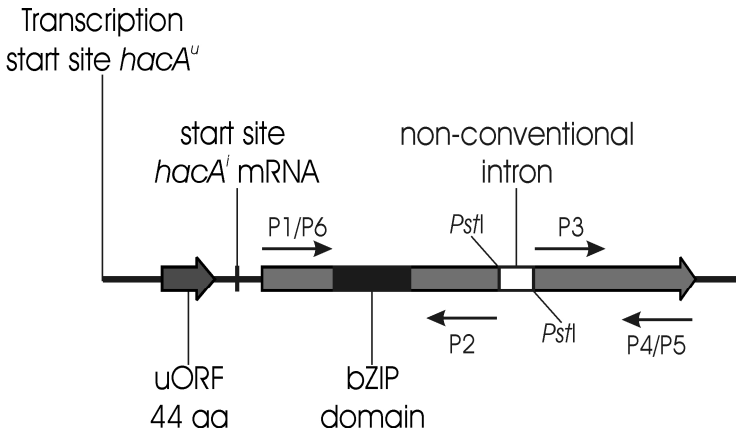


Figure 1. Schematic representation of *hacA*^{u} mRNA with its main features. The locations of oligonucleotide primers used for the construction pHM50, pHM52Δ4, and pHM30Δ3 are indicated by the small arrows.

For over-expression in *A. niger* the 5' and 3' portions of the *hacA* ORF were amplified using the primer sets P1-P2 and P3-P4, respectively. The two fragments were then fused at the *PstI* site and cloned into pAN52-1Not. *A. nige*r N592 was transformed with the resulting plasmid (pHM50). For expression in *E. coli* a construct was generated in the same way, except that primer P4 was replaced by primer P5, which codes for a $(His)_6$ tag at the C-terminus of the HacA protein. This construct was cloned into pET3d+ (Novagen), and the resulting plasmid (pHM30Δ3) was transformed into the *E. coli* strain BL21 (DE3)pLysS (Novagen). HacA-$(His)_6$ was purified by FPLC (Amersham Pharmacia Biotech), using a Ni-NTA Superflow column (Qiagen), followed by gel filtration on a Superdex 75 column (Amersham Pharmacia Biotech). For construction of a *gfp-hacA*^{i} fusion, *hacA*^{i} was amplified with the primers P4 and P6 using pHM50 as template, and *gfp* was generated with primers P7 and P8 using pEGFP (Clontech) as template. The two genes were ligated together at the *NarI* site, and cloned into pAN52-1Not. Subsequent ligation of the *pyrA* gene into this plasmid resulted in pHM52Δ4, which was used to transform *A. niger* N592. Maps of the plasmids used in this study are shown in Figure 2.

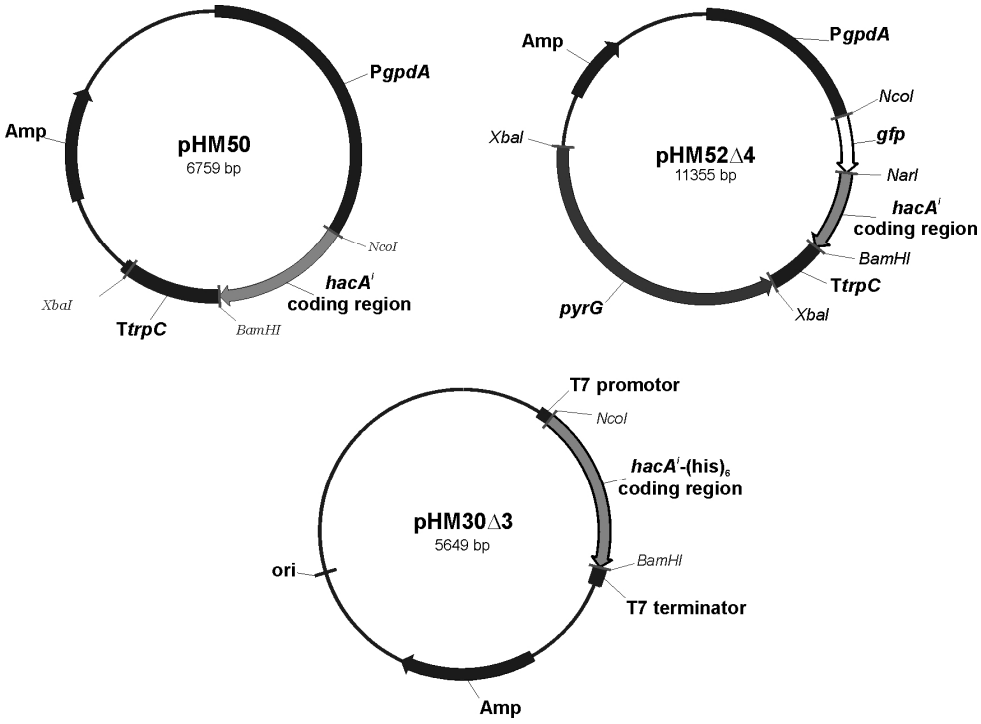


Figure 2. Maps of the plasmids used in this study.

Northern Blot Analysis

For RNA isolations, mycelium was ground with 1mm glass beads using the Fast Prep FP120 (BIO 101 Savant), and total RNA was isolated using the RNeasy plant total RNA kit (Qiagen). For Northern analysis 10 µg aliquots of total RNA were fractionated on a 1.5% (w/v) agarose-formaldehyde gel and blotted onto a Hybond N membrane (Amersham Pharmacia Biotech). To generate probes, DNA fragments were α-^{32}P-labelled using the Prime-It II labelling kit (Stratagene), and the probes were purified on Biogel P30 spin columns (Biorad). Northern hybridisations were performed at 42°C in 50% formamide containing 10% polyethylene glycol 6000, 5x Denhardt's solution, 50 mM TRIS-HCl (pH 7.5), 1 M NaCl, 0.1% sodium pyrophosphate, 1% SDS and 100 µg/ml salmon sperm DNA (Sigma). Filters were washed 20 minutes with 0.2 x SSC / 0.1% SDS at 60°C.

Western Blot Analysis

Criterion SDS-polyacrylamide (PAGE) gels (Biorad) were used for protein electrophoresis. SDS-PAGE gels were stained with Coomassie Brilliant Blue R250 (Biorad) or were blotted onto pure nitrocellulose membranes (Biorad) for Western blot analysis. Protein concentrations were determined according to Bradford [9] using BSA as a standard. Western analysis was performed using standard procedures [141]. Membranes were incubated in 100 ml TN (50 mM TRIS-HCl, 150 mM NaCl, pH 8.0) containing 5 g low fat milk powder and polyclonal antibodies specific for BIPA, CYPB, and PDIA. Immunoreactive proteins were detected using an HRP-conjugated secondary antibody and ECL detection reagents (Amersham Pharmacia Biotech).

Southern Blot Analysis

Genomic DNA for Southern analysis was isolated by adding extraction buffer (100 mM TRIS-HCl pH 8.0, 50 mM EDTA, 500 mM NaCl, 1mM DTT, 35 mM SDS) to ground mycelium. After 10-min incubation at 65 °C, mycelium debris was separated by centrifugation, and the DNA in the supernatant was precipitated with isopropanol. The DNA pellet was resuspended in H_2O and incubated with DNase-free RNase (Roche Molecular Biochemicals). For each digest 10 µg of genomic DNA was incubated over night at 37 °C with either 20 U of either *NotI* or *NcoI* endonuclease. The reactions were fractionated on a 0.8% TAE agarose gel. Blotting and hybridisation were done according to standard procedures [141].

5'RACE

5'RACE (Rapid Amplification of cDNA Ends) was performed on 1 µg of total RNA using the SMART RACE cDNA Amplification Kit (Clontech). Primer P9 (Table 1) was used to amplify *hacA* cDNA.

Electrophoretic Mobility Shift Assay (EMSA)

The promoter fragments of *bipA*, *cypB*, *pdiA*, *prpA*, and *tigA* (Table 2) were synthesised by PCR performed on gDNA of *A. niger*. The resulting fragments were gel purified and end-labelled with γ-^{32}ATP using T4 polynucleotide kinase (New England Biolabs) and purified on Biogel P30 spin columns (BioRad). Binding reactions (40 µl) included 0.5 µg of purified HacA protein, and 1x binding buffer (20 mM HEPES pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 0.25 mM EDTA, 0.5 mM DTT, 2% Ficoll 200, 100 µg/ml poly(dI-dC), and 5% glycerol). 1 ng of radiolabelled probe was added and binding was allowed to proceed on ice for 15 min. The reactions were loaded on a 5% polyacrylamide gel and electrophoresed at 100 V for 3 hr in 0.25 x TBE.

Table 2. Promoter fragments used in EMSAs

Gene	**Accesion No.**	**Position relative to ATG**
bipA	Y08868	-412 to 0
cypB	AY005867	-396 to 0
pdiA	X98797	-432 to 0
prpA	AF095899	-587 to 0
tigA	X98748	-330 to 0
hacA	AY303684	-863 to -265

UV Microscopy

Transformants expressing GFP-HacA were grown in minimal medium as described above. Mycelium samples were examined by fluorescent microscopy on a Polyvar microscope (Reichert-Jung) coupled to a cooled CCD digital camera, using FITC filters. Pictures were analysed with the Metamorph software package (version 4.5). DAPI (Molecular Probes) was used at a concentration of 1 µM for visualisation of nuclei.

Results

Cloning and Sequence Analysis

The gene encoding the transcriptional activator of the unfolded protein response in *A. nidulans*, *hacA* , has previously been cloned based on its homology to the *HAC1* gene of *S. cerevisiae* [140]. The *A. nidulans hacA* sequence was used in this study as a probe to facilitate cloning of the corresponding gene from a cDNA library derived from the *A. niger* var. *awamori* strain UVK143f. Several positive clones were obtained, and sequencing of their 5' ends revealed that they were indeed derived from a homologue of the *A. nidulans hacA* gene. Based on the DNA sequence, primers were designed, and a 6.2-kb genomic fragment including the *hacA* coding sequence together with its flanking regions (Accession No. AY303684) was cloned from *A. niger* N402 by genomic walking. Sequence comparison of the *hacA* cDNA from *A. niger* var. *awamori* with the *hacA* gene from *A. niger* N402 revealed that the two encode identical proteins.

In *S. cerevisiae*, an unconventional 252-bp intron is spliced from the *HAC1* mRNA during UPR, whereas a much shorter (26-bp) intron is excised from the mammalian *XBP1* mRNA under similar conditions [12,22]. The *hac1*/*hacA* genes from *T. reesei* and *A. nidulans* have also been shown to contain non-conventional introns that are spliced during the induction of UPR [140]. Comparison of the cDNA sequence of one of the clones obtained from the cDNA library of *A. niger* var. *awamori* strain UVK143f with the corresponding genomic

sequence revealed a non-conventional intron of 20 nt. The splicing site of the intron could not be determined unambiguously by comparing the unspliced and spliced sequences, because of the presence of a CTGCAG sequence on both sides of the intron. However, the intron borders and splice-sites of the yeast *HAC1* and mammalian *XBP1* are well conserved [12], and these conserved bases were also found at the borders of the *A. niger hacA* (Fig. 3). Moreover, a strong secondary RNA structure was predicted for the intron region and its flanking sequences. The structure of the conserved intron borders—organised in two loops of 7 nt each—was similar to those formed by both the yeast *HAC1* and mammalian *XBP1* introns. Most probably, the *A. niger hacA* transcript is also spliced within these loops.

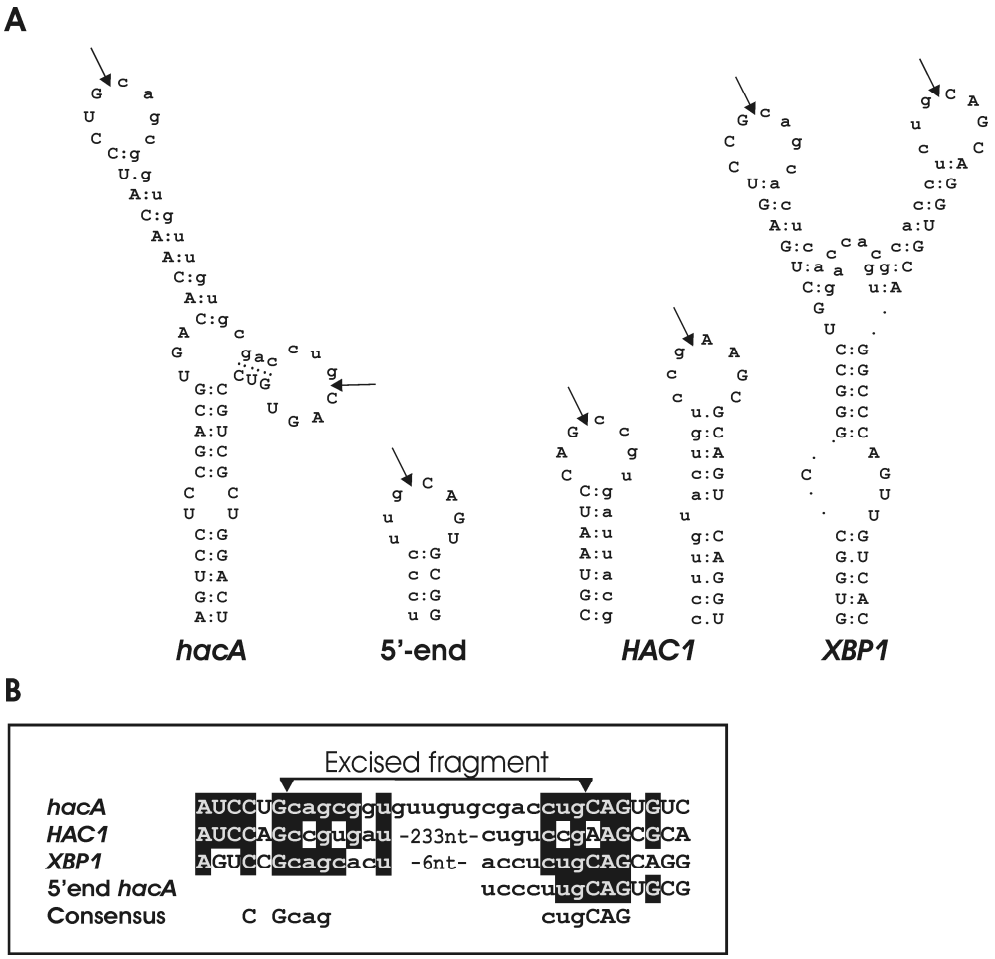


Figure 3. Structure and context of the *hacA* unconventional intron. **(A)** Predicted secondary structure of the *A. niger hacA* intron, the 5'-end of the *hacA*^{i} mRNA, the 5' and 3' border of the *S. cerevisiae HAC1* intron, and the mammalian *XBP1* intron, respectively at the RNA level. The intron sequences are shown in lower case letters, and the arrows indicate the cleavage sites. **(B)** Alignment of the mRNA sequence surrounding the unconventional intron in *A. niger hacA*, *S. cerevisiae HAC1*, and the mammalian *XBP1*. The 5'-end of *A. niger hacA*^{i} mRNA is also indicated. Conserved residues in the loops are indicated in the consensus sequence.

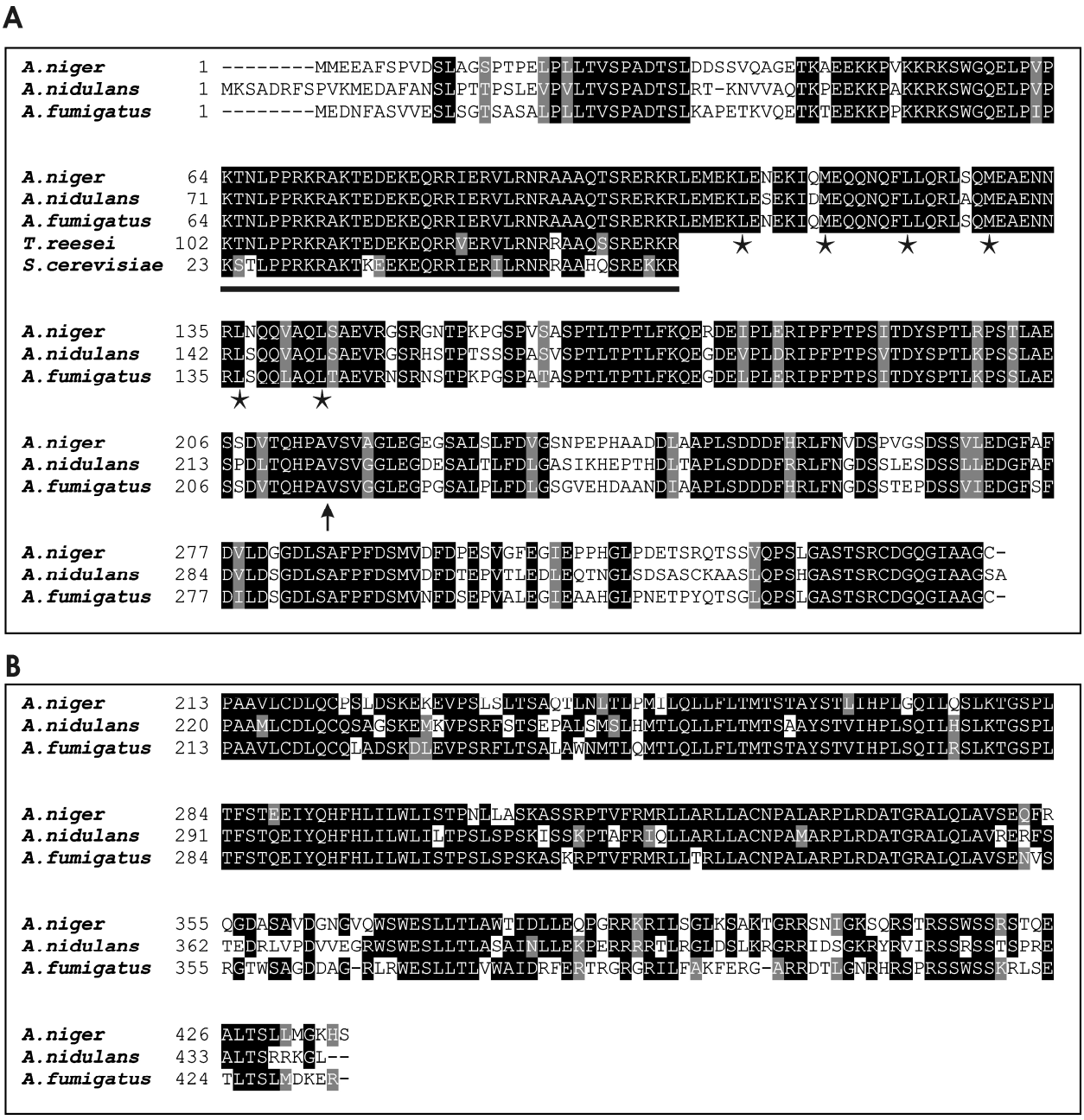


Figure 4. Alignment of the amino acid sequences of the induced and uninduced forms of HacA from *A. niger*, *A. nidulans* and *A. fumigatus*. Only the DNA binding domains of the *S. cerevisiae* Hac1p and the *T. reesei* Hac1 proteins are shown. **(A)** The putative DNA binding domain is underlined the residues of the putative leucine zipper are indicated by the *asterisks*, and the position of the unconventional intron in the three *Aspergillus* species is indicated by the arrow. **(B)** Alignment of the uninduced forms of HacA from *A. niger*, *A. nidulans* and *A. fumigatus* from the site of the unconventional intron to the C-terminus.

The unspliced *hacA*^{u} mRNA contains an ORF coding for a 436-amino acid protein. Splicing replaces a C-terminal portion coding for 222 amino acids with a different sequence coding for 128 amino acids. This differs from the situation in yeast, where a 10-residue C-terminal tail is replaced by a sequence of 18 amino acids as a result of splicing [22].

The ORF of the *A. niger hacA*^{i} mRNA encodes a protein of 342 amino acids, which is 73% and 80% identical to its orthologues from *A. nidulans* and *A. fumigatus*, respectively (Fig. 4a). HacA belongs to the bZIP type family of transcription factors, which typically have a leucine-zipper dimerisation motif downstream of the DNA-binding domain. In *A. niger* HacA there are six amino acids (L, M, L, M, L and L) that could potentially form the leucine zipper.

Interestingly, alignment of the protein sequences encoded by the unspliced *hacA*^{u} mRNAs from *A. niger*, *A. nidulans* and *A. fumigatus* showed a high degree of identity (71%–75%) among the species, even in the part downstream of the non-conventional intron (Fig. 4b). This could indicate that the HacA protein encoded by *hacA*^{u} mRNA also has a function for which the C-terminal segment is important.

Northern Analysis of ER Stress

One of the key steps in activation of the UPR pathway in yeast and mammals is the Ire1p/IRE1 mediated splicing of a unconventional intron from the mRNA encoding the transcriptional activator of the UPR [12,22]. A similar event has been shown to occur during the induction of the UPR in *T. reesei* and *A. nidulans* [140]. To study this event in *A. niger*, ER stress was triggered by three means: (1) over-expression of a heterologous protein (t-PA), (2) treatment of mycelium with DTT, and (3) treatment of mycelium with tunicamycin. DTT and tunicamycin are both known to be chemical inducers of the UPR. Tunicamycin provokes the UPR by inhibiting the core oligosaccharide addition to nascent polypeptides in the ER, thereby blocking proper folding and transit through the ER [122]. As a reducing agent, DTT disrupts the oxidising environment of the ER, preventing the formation of disulphide bonds. Tissue plasminogen activator (t-PA) is a serine protease which, when activated by fibrin, itself activates plasminogen, enabling the dissolution of blood clots. Low expression levels have hampered efforts to produce useful amounts of t-PA in various host organisms, and in *A. niger* it has been used as a model protein to assess the problems of heterologous protein expression [181]. Since expression of t-PA in *A. niger* leads to the transcriptional upregulation of genes encoding the ER chaperones and foldases BIPA, CYPB, and PDIA, it was chosen as a model strain to assess the UPR under conditions of heterologous protein production.

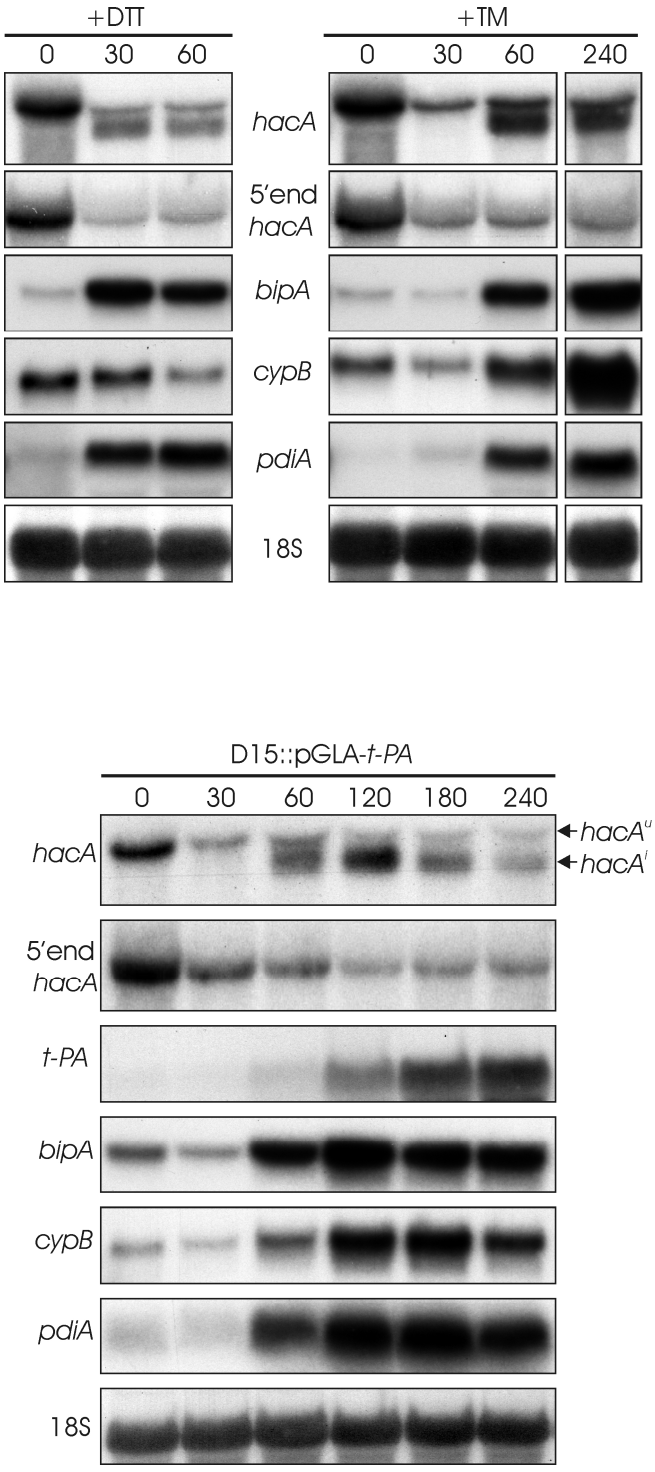


Figure 5. Northern hybridization analyses showing the effects of ER stress, imposed by treatment of cells with DTT or tunicamycin (TM) and by the expression of a heterologous gene (D15::*pGLA*-t-PA), on the expression of *hacA*, *bipA*, *cypB*, and *pdiA*. The time in minutes after induction is indicated above the lanes. The two *hacA* mRNA species that were present after induction of ER stress are indicated by the arrows, and are labelled $hacA^u$ and $hacA^i$, for uninduced and induced, respectively. The gene coding for 18S rRNA was used as a control probe.

To examine the UPR induced by the synthesis of t-PA, or by treatment with DTT or tunicamycin, Northern hybridisation was performed on total RNA isolated from mycelia of such cultures. A cDNA fragment comprising the region from the start codon up to the unconventional intron was used as a probe to detect *hacA* mRNA. Northern analysis showed that a second, smaller *hacA*^{i} mRNA transcript emerged within 30 minutes after exposure of the mycelium to DTT, and this coincided with a decrease in the intensity of the *hacA*^{u} band (Fig. 5). Together with the appearance of the smaller *hacA*^{i} mRNA, upregulation of the ER chaperone-encoding gene *bipA* [168], and the foldases-encoding gene *pdiA* [104] was observed. Although some fluctuations in the mRNA levels for the ER foldases CypB [27] were observed, its expression was not obviously affected by the DTT treatment. In contrast, tunicamycin treatment had a profound effect on the expression of *cypB*, and led to a seven-fold upregulation within 4 h (phosphor-imager data, results not shown). Like DTT treatment, tunicamycin treatment also led to the appearance of the *hacA*^{i} band, and the simultaneous upregulation of *bipA* and *pdiA*, although the effect appeared later than in the DTT treated mycelia. Northern analysis of *A. niger* D15::Pgla-t-PA#19 showed that the effect on the UPR was already manifested by in the up-regulation of *bipA*, *cypB* and *pdiA*, and the formation of a smaller *hacA*^{i} band, before clear induction of t-PA could be detected (after 2 h of induction), underscoring the sensitivity of the UPR. Northern blot analysis also revealed that the transcript levels of *bipA*, *cypB* and *pdiA* increased in parallel with the increase in the t-PA transcript. From these results it appears that the effects of ER stress, imposed by the expression of a heterologous protein more closely resembles that caused by tunicamycin than that by treatment with DTT. This is valid for both the response time, and the transcription levels of *bipA*, *cypB* and *pdiA*. The total amount of the *hacA* transcript is lower at t_{30} compared to t_0 in all three ER stress treatments. This could be the result of transcriptional down-regulation due to the transfer of mycelia at the start of the experiments. Another possible explanation is that the processed *hacA*^{i} mRNA is less stable than the uninduced *hacA*^{u}. This could give the fungus the ability to shut down the UPR more rapidly when the ER stress has been relieved.

5'-Truncation of *hacA* mRNA

Splicing of the 20-nt non-conventional intron upon UPR induction was demonstrated by comparative analysis of the cDNA and genomic sequences. However, this splicing event alone cannot account for the difference in size between the *hacA*^{u} and *hacA*^{i} transcripts observed by Northern analysis. Sequence analysis of the clones obtained from the *A. niger* cDNA library indicated that *hacA*^{i} might also be truncated at the 5'-end relative to *hacA*^{u}. Furthermore, the *A. nidulans* and *T. reesei hacA/hac1* mRNAs are cleaved within the 5'-flanking region upon UPR induction. However, neither the mechanism nor the exact site of cleavage has been determined in those cases [140]. To confirm that truncation of the 5'-end of

hacA mRNA takes place during UPR in *A. niger*, Northern blots were probed with a 230-bp fragment comprising the most upstream part of the *hacA* 5'UTR (Fig. 5). As expected, this probe only hybridises to the full-length $hacA^u$ mRNA, and the signal decreases upon induction of the UPR. Furthermore, this probe does not hybridise to the $hacA^i$ mRNA. This confirms that the $hacA^i$ transcript also lacks a 5'-segment of the $hacA^u$ mRNA. To determine the site of truncation, and to confirm the splicing of the 20-nt unconventional intron from the *A. niger hacA* mRNA under UPR conditions, a 5'RACE experiment was performed. The reactions were performed on RNA extracted from *A. niger* D15::pglaA-t-PA before, and 4 h after UPR induction (Fig. 5). A gene-specific primer, located downstream of the unconventional intron, was used to amplify the *hacA* cDNA. A single band was obtained from a mycelial sample harvested before induction. Two bands were obtained from the sample harvested after induction; the larger of these two bands had the same size as the fragment obtained before induction. This is in agreement with the Northern blot data, which indicated that two *hacA* mRNA species ($hacA^u$ and $hacA^i$) were present 4 h after induction (Fig. 5). All the fragments obtained were cloned and sequenced. The fragment derived from the sample taken before induction was identical to the larger fragment from the sample taken 4 h after t-PA induction ($hacA^u$), and contained the unconventional 20-nt intron. The 5'-ends were located 304 nt upstream of the *hacA* start codon (Fig. 6).

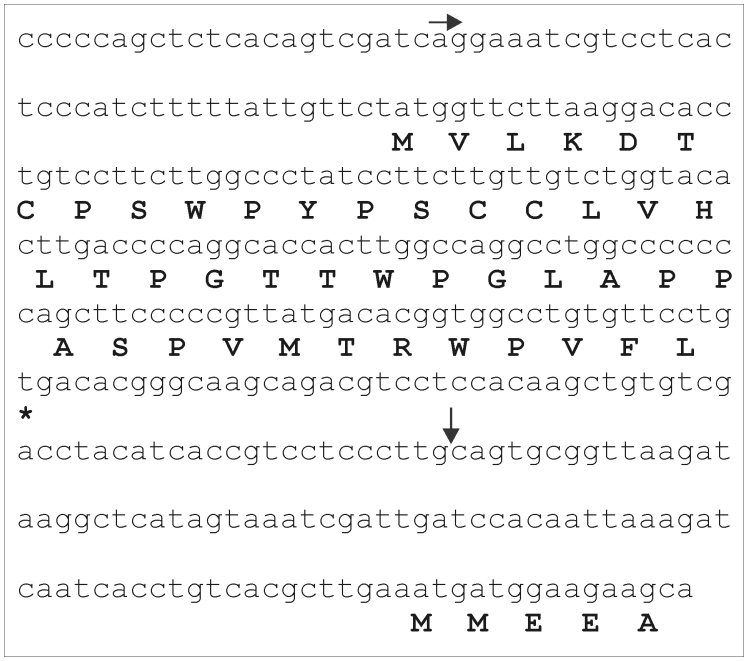


Figure 6. Overview of the sequence at the 5'UTR of the *A. niger hacA* gene. The first five amino acids of HacA are indicated in upper case letters. The transcription start site of $hacA^u$ is indicated by the horizontal arrow, and the start site of $hacA^i$ by the vertical arrow. The ORF encoding a 44-amino acid peptide is also indicated in upper case.

Sequence analysis of the smaller fragment, obtained from the induced sample (*hacA*^{i}), revealed that its 5'-end lies 230 nt further downstream, only 74 nt upstream of the *hacA* start codon (Fig. 6). Furthermore, the unconventional 20-nt intron was not present in these fragments, confirming that splicing had taken place. The part removed from the 5'UTR of *hacA* contains a short ORF encoding a 44-amino acid peptide. Although it is not known whether this extra region is functional, it could imply the involvement of a translational control mechanism. Interestingly, further inspection of the sequence around the truncation site revealed a sequence stretch similar to those found around the sequence borders of the 20-nt intron (Fig. 3).

Over-expression of *hacA*

To determine whether over-expression of *hacA*^{i} would mimic the effects of t-PA expression, DTT treatment or tunicamycin treatment, the *hacA* coding region was cloned in an *Aspergillus* expression vector, under the control of the constitutive *A. nidulans gpdA* promoter, and the resulting plasmid (pHM50) was introduced into the genome of *A. niger* N592.

Southern analysis of two transformants detected two additional copies of the *hacA* gene in N592::pHM50#1, and one additional copy in N592::pHM50#2, relative to wild type (Fig. 7). Northern analysis of these strains showed substantial accumulation of *bipA*, *cypB*, and *pdiA* mRNAs in N592::pHM50#1 (Fig. 8a), and a minor up-regulation in the single-copy integrant N592::pHM50#2. These results were confirmed by Western analysis (Fig. 8b), and are analogous to the results obtained upon treatment with tunicamycin and induction of t-PA expression. Quantification of the *hacA*^{u} mRNA by phosphor-imager analysis (data not shown) revealed four- and two-fold higher signals in N592::pHM50#1 and N592::pHM50#2, respectively, compared to wild type. This up-regulation of *hacA*^{u} upon over-expression of *hacA*^{i} suggests that, under UPR conditions, transcription of the *hacA* gene is up-regulated by its own gene product.

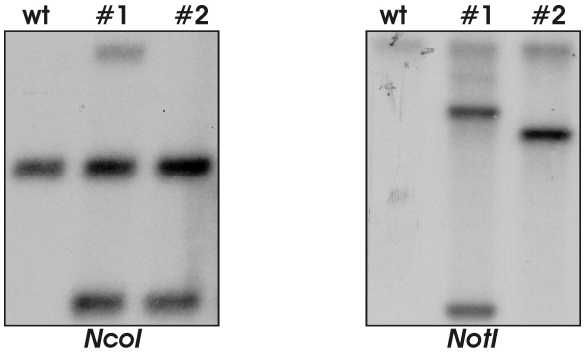


Figure 7. Southern analysis showing two additional copies of the *A. niger hacA* gene in strain N592::pHM50#1, and one additional copy in N592::pHM50#2, compared to the wild-type strain N592.

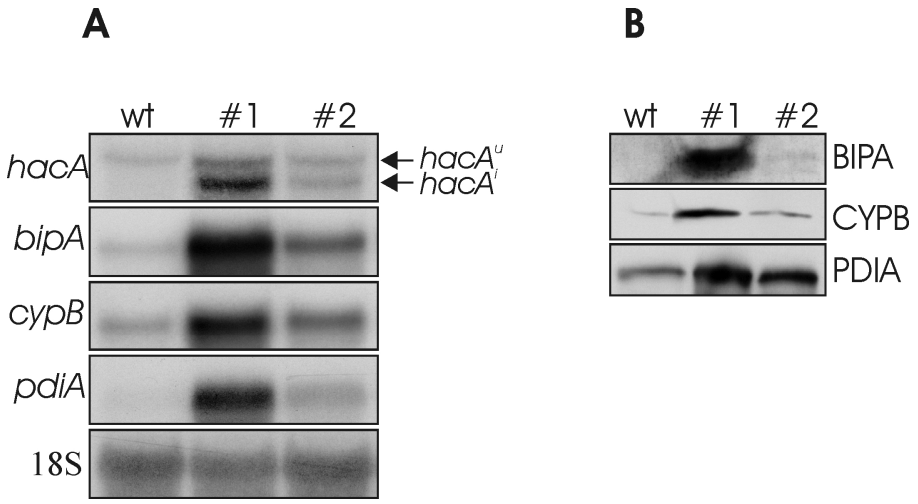


Figure 8. Effects of over-expression of *hacA*^{i} on the expression of *bipA*, *cypB*, and *pdiA*. **(A)** Northern analysis showing the effect of over-expression of *hacA*^{i} in *A. niger*. Probes used in hybridisations are indicated on the left. (Lanes: wt, wild-type *A. niger* strain N592; #1, *A. niger* strain N592::pHM50#1; #2, *A. niger* strain N592::pHM50#2). The arrows indicate the two *hacA* mRNA species *hacA*^{u} and *hacA*^{i}. Note that the exposure time for the *hacA* blot was reduced relative to that used for Fig. 5, due to the high expression level in strain #2. **(B)** Western analysis showing the effect of *hacA*^{i} over-expression on the levels of BipA, CypB and PdiA protein. Antibodies are indicated in upper case letters on the right.

Electrophoretic Mobility Shift Analysis (EMSA)

The experiments described so far show that up-regulation of ER resident chaperones and foldases occurs synchronously with processing of *hacA* mRNA during ER stress. Similar results were obtained without secretion stress upon over-expression of *hacA*^{i}, suggesting that HacA is the factor responsible. Whether HacA indeed binds to the promoters of these up-regulated genes was studied *in vitro* by EMSA. Purified HacA protein was used to evaluate binding to the promoters of *bipA*, *cypB*, *pdiA* and *hacA* itself, and to the promoters of two additional ER foldase-encoding genes, *prpA* and *tigA*. *prpA* is homologous to the protein disulfide isomerase gene *pdiA*, and it was shown to be up-regulated during heterologous protein biosynthesis [174]. The product of *tigA* also belongs to the PDI family, and its transcription is induced upon treatment with tunicamycin [62]. The promoter of *cypA*, which encodes a cytoplasmic cyclophylin [26], was used as a negative control. Band-shifts were observed for all the ER-specific promoters (Fig. 9). Interestingly, binding was also observed with a *hacA* promoter fragment. This is in accordance with the finding that over-expression of *hacA*^{i} leads to up-regulation of *hacA*^{u}.

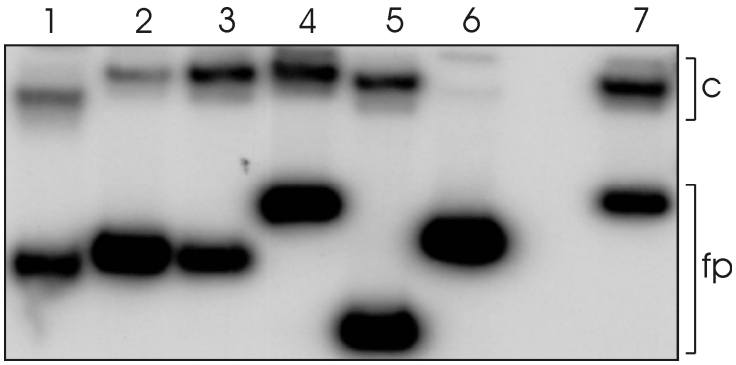


Figure 9. Electrophoretic mobility shift analysis (EMSA) showing the binding of the *A. niger* HacA protein to the promoters of genes encoding ER chaperones and foldases, and to the promoter of *hacA* itself (c, protein-DNA complex; fp, free probe). Lanes: 1, P*bipA*; 2 P*cypB*; 3, P*pdiA*; 4, P*prpA*; 5, P*tigA*; 6, P*cypA* (negative control); 7, P*hacA*. The sizes of the promoter fragments used can be deduced from the data in Table 2.

Cellular Localisation of HacA

Analysis of HacA using the PSORT II program (http:// psort.nibb.ac.jp/) identified two putative Nuclear Localisation Signals (NLSs), a putative monopartite NLS (PVKKRKS comprising amino acid residues 49– 55), and a putative bipartite NLS (RKRAKTEDEKEQRRIER, amino acid residues 70– 86). The putative bipartite NLS is located in the DNA-binding domain. To examine the cellular localisation of HacA, a hybrid *gfp-hacA* construct was made, consisting of the *hacA*^{i} ORF fused to the 3'-end of the *gfp* ORF. Fluorescence microscopy of N592::pHM52D4, which expresses *gfp-hacA*^{i} under control of the *gpdA* promoter, showed a similar pattern of fluorescence as hyphae stained with the nuclear dye DAPI, confirming that HacA indeed localises to the nucleus (Fig. 10).

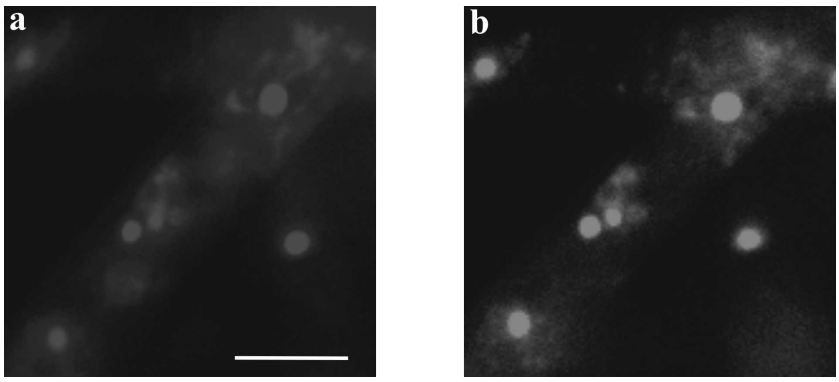


Figure 10. Fluorescence microscopy showing the localisation of GFP-HacA. **(A)** Staining of the hyphae with the nuclear stain DAPI. **(B)** Localisation of GFP. The bar represents 10 µm.

Discussion

The *A. niger hacA* gene was isolated based on its similarity to the *A. nidulans hacA* sequence. The HacA proteins from *A. niger* and *A. nidulans* are very well conserved and are 73% identical. The putative HacA from *A. fumigatus* found in the TIGR genome database (http://www.tigr.org/tdb/e2k1/afu1/) is 80% identical to the *A. niger* HacA. The DNA-binding domains of the three HacA homologues are identical, while their leucine zipper domains are very similar to each other. Sequence differences are mostly located in the N-terminal and C-terminal parts of the proteins.

In yeast, *HAC1* mRNA is expressed relatively abundantly in unstressed cells, but it is not translated due to the presence of a 252-nt intron. This intron pairs with the 5'UTR, thereby preventing the read-through by the ribosomes [137]. Splicing of this intron by Ire1p is necessary for UPR induction [22].

Although the mammalian XBP1 is similar to the yeast Hac1p only insofar as both are bZIP proteins, it has been shown that nearly all features of the yeast splicing pattern are conserved in mammals [186]. However, in contrast to *HAC1*, which is transcribed constitutively, *XBP1* is expressed at low levels and the mRNA is translated, since no inhibitory secondary structure can be formed. Furthermore, activation of the mammalian UPR involves two transcriptional activators, ATF6 and XBP1. Activation starts with the proteolytic release of ATF6 from the ER. ATF6 then up-regulates transcription of genes for ER chaperones, as well as the expression of *XBP1*. *XBP1* is activated by IRE1-mediated splicing of a 26-nt intron, and its protein product up-regulates the same targets as ATF6, including *XBP1* itself. It has been suggested that this system would provide the mammalian cell with a more flexible UPR which is capable of coping more effectively with ER stress [186]. The similarity between the *A. niger* HacA and the yeast Hac1p is confined to the DNA binding domain, and no significant homology could be found between the *Aspergillus* protein and its mammalian orthologue XBP1. Despite these differences, our analysis of the regulation of the *A. niger* HacA indicates that it shares features of both the yeast and the mammalian equivalents, and might even possess features that are specific for the UPR of *A. niger* and other filamentous fungi.

The unconventional introns of *HAC1* mRNA and *XBP1* mRNA are spliced by Ire1p/IRE1 upon induction of the UPR [152,186]. The unconventional 20-nt intron present in the *A. niger hacA*^{u} shows a close resemblance to its counterparts in yeast and mammalian cells (Fig. 3) [12,152]. The sequences at the boundaries of the intron are well conserved, and a similar secondary RNA structure is predicted where the intron borders are encompassed in two loops of 7 nt. This structural and sequence similarity suggests that the *A. niger hacA*^{u} RNA is spliced in a similar manner, most probably by the action of an Ire1p/IRE1 like

protein. We have recently cloned a gene from *A. niger* with similarity to the yeast Ire1p, and preliminary results suggest it is involved in the splicing of *hacA* mRNA.

The only known modification of the yeast *HAC1* and mammalian *XBP1* mRNAs that is triggered by UPR induction is the splicing of the unconventional intron. Results from *A. nidulans* and *T. reesei* [140], and data presented in this paper for *A. niger*, indicate that the *hacA*^{u} mRNA undergoes an additional modification in these fungi. In addition to the unconventional intron-splicing event induced by ER stress, the 5'UTR of *hacA* mRNA is truncated. A 230-nt segment is removed, which contains an ORF with the capacity to code for 44 amino acids. The 5'UTR of the *A. nidulans hacA* gene contains an uORF for seven amino acids, and the *T. reesei hac1* gene contains one uORF for 18 amino acids and one for two amino acids [140]. The presence of these uORFs suggests the involvement of a translational control mechanism. Translational control is often associated with short upstream ORFs in the 5'UTRs of mRNAs [90]. Indeed translation of the yeast transcription factor Gcn4p is controlled by four uORFs [59], and translational control of the synthesis of the transcription factor ATF4, which is involved in the mammalian stress responsive pathway, is dependent on two uORFs [51]. The mechanism responsible for the truncation of *A. niger hacA* mRNA upon ER stress is still unknown. Our results revealed sequence similarity between the truncation site in the 5'-portion of *hacA* mRNA and the borders of the unconventional intron (Fig. 3). However, no RNA secondary structure incorporating the putative truncation site in a stable stem-loop could be predicted for the region around the 5'-truncation site. This makes an involvement of IreA in the truncation unlikely. Furthermore, it has been speculated that, upon induction of the UPR, transcription of *hac1* in *T. reesei* might initiate at a new start site, thus giving rise to the truncated *hac1* mRNA [140]. Our recent identification of a *cis*-acting UPR element in the 5'UTR of the *A. niger hacA* gene supports this hypothesis [100].

A. niger HacA is most probably involved in the upregulation of its own gene, as purified HacA protein was able to bind to the *hacA* promoter (Fig. 9) Moreover, over-expression of *hacA*^{i} not only led to up-regulation of *bipA*, *cypB* and *pdiA*, but also led to the accumulation of *hacA*^{u} mRNA (Fig. 8). This is in agreement with the finding that in mammalian cells the transcription of *XBP1* is up-regulated by XBP1 protein upon the onset of the UPR [186].

In conclusion, the several lines of evidence presented here strongly indicate that the *hacA* gene in *A. niger* encodes a transcription factor that regulates the UPR.

Chapter 3

HacA, the Transcriptional Activator of the Unfolded Protein Response in *Aspergillus niger*, Binds to Partly Palindromic UPR Elements of the Consensus Sequence 5'-CAN(G/A)NTGT/GCCT-3'

Harm J. Mulder, Igor Nikolaev, Susan M. Madrid

Published in Fungal Genetics and Biology. 2006, 43: 560-572

Abstract

The promoters of UPR target genes contain an unfolded protein response element (UPRE), which confers the stress inducibility to the gene, via an interaction with the transcription activator HACA. In the promoters of the *Aspergillus* ER-stress responsive genes *bipA, cypB, pdiA, prpA, tigA, and hacA,* a consensus sequence was identified, which was located close to the transcription start site of the gene (< 81 bp), and corresponds to the sequence CAN(G/A)NTG**T/GCCT**. The UPRE is a partly palindromic sequence around a dispensable spacer nucleotide, followed by four highly conserved bases. By an *in vitro* selection procedure, an optimal binding site for HacA was isolated. This sequence, ACACGTGTCCT, resembles the UPRE but lacks the spacer nucleotide. It has a much higher binding affinity than the identified UPREs, and *in vivo* it behaves as a more powerful *cis*-acting element.

Introduction

Filamentous fungi, such as *Aspergillus* and *Trichoderma* species are important industrial production organisms due to their ability to secrete large amounts of native proteins. Expression of heterologous proteins is however often hampered by low expression yields [47]. Although several strategies such as gene fusions, the development of protease deficient host strains, and co-over-expression of foldases have sometimes improved the yields of some heterologous proteins, in other cases no improvement was found [18]. A more detailed insight in the fungal secretory pathway may lead to the identification of bottlenecks and problems in heterologous protein production.

In eukaryotic cells, newly synthesised secretory and transmembrane proteins are folded and assembled in the endoplasmic reticulum (ER). The ER provides an oxidising environment in which protein folding is assisted by a number of molecular chaperones and folding enzymes [39,56]. Protein folding in the ER can be compromised by several endogenous and exogenous factors like changing environmental conditions, cell differentiation, chemical insult, pathogenic infections, and expression of heterologous proteins [138]. This ultimately results in the accumulation of unfolded proteins within the ER, since only correctly folded proteins transit the secretory pathway [28]. The eukaryotic cell reacts to the accumulation of unfolded proteins in the ER by inducing a pathway known as the unfolded protein response (UPR) [64,94,150,179]. The UPR is a versatile regulatory system that serves to maintain homeostasis of ER functions under stress conditions.

The UPR pathway has been well characterised in *Saccharomyces cerevisiae*. Three key components of this signal transduction pathway in *S. cerevisiae* are: (1) Ire1p, a serine/threonine protein kinase with an RNaseL like endonuclease domain. Ire1p resides in the ER membrane, and was identified as the proximal sensor of the UPR [21,96]. (2) Hac1p, a basic leucine zipper (bZIP) type of transcription factor that up-regulates the transcription of various target-genes of the UPR pathway [22,95]. (3) The unfolded protein response element (UPRE), a *cis*-acting element present in the promoters of UPR-target genes, and a target for Hac1p [97].

Ire1p senses the accumulation of unfolded proteins through a dynamic interaction with the molecular chaperone BiP [8,115]. Release of BiP triggers the dimerisation and subsequent trans-phosphorylation of Ire1p, thereby activating its RNaseL like endonuclease domain. This ultimately results in the activation of *HAC1* mRNA, by the splicing of an unconventional intron from the messenger [43,152]. The spliced *HAC1* mRNA is efficiently translated, resulting in the mature form of Hac1p, the transcriptional activator of the UPR [14,66,137].

Although the pathway is much more elaborate in higher eukaryotes, the central activation step of the transcriptional activator of the UPR by unconventional splicing of its mRNA has been conserved [191]. Also in the filamentous fungi *Aspergillus niger*, *Aspergillus nidulans* and *Trichoderma reesei*, activation of the UPR involves unconventional intron splicing from the *hacA/hac1* mRNA. However, in addition to the unconventional intron splicing, the 5'-untranslated region (5'UTR) of the transcript is truncated upon ER stress [101,140]. This characteristic has not been reported for the *S. cerevisiae HAC1* and mammalian *XBP1* homologues [12,67]. The mechanism behind the truncation of the 5'UTR of filamentous fungal *hacA/hac1* mRNA is still unknown, but this could offer the fungus an extra control point in the activation of the UPR [101,140].

Analysis of the promoter regions of several UPR target genes has revealed different *cis*-acting elements in *S. cerevisiae* and mammalian cells. In *S. cerevisiae*, the *KAR2* UPRE corresponds to CAGCGTG, a partly palindromic sequence separated by a spacer nucleotide [97]. In mammals, several different *cis*-acting elements were identified. An UPRE containing the consensus sequence TGACGTGG/A, is recognised by ATF6 [176]. Two ER stress response elements, ERSE1 corresponding to CCAAT-N_9-CCACG [136,185], and ERSE2 corresponding to ATTGG-N_9-CCACG [74] are targets for both ATF6 and XBP1. In *Aspergilli,* an UPRE-like consensus sequence has not been characterised thus far, but several UPR responsive genes have been cloned, mostly encoding ER localised protein foldases. The transcriptional behaviour of these genes upon ER stress indicates the presence of an UPRE in their promoters. Among them are the *A. niger* genes *bipA*, encoding a molecular chaperone, *cypB* encoding a prolyl-*cis*-*trans*-isomerase, and three genes encoding the protein disulphide isomerases *pdiA*, *prpA* and *tigA* [27,62,104,168,174].

Genomic analyses in *S. cerevisiae, Arabidopsis thaliana* and *A. nidulans* have linked a large number of genes to the UPR, affecting virtually every stage of the secretory pathway [89,154,160]. The breadth of the UPR emphasises the necessity to understand the pathway in detail, in order to be able to optimise the fungal secretory pathway for heterologous protein production.

In this paper, we describe the identification and characterisation of *Aspergillus* UPREs in the promoters of five ER resident chaperones/foldases and in the promoter of *hacA* itself. Furthermore, we demonstrated by mutation analysis the importance of the leucine zipper domain of HacA in binding the UPREs.

Material and Methods

Strains and Transformations

A. niger AB4.1 (*cspA1, pyrG*) was grown on minimal media plates containing per litre: 6 g $NaNO_3$, 1.5 g KH_2PO_4, 0.5 g $MgSO_4 \cdot 7\ H_2O$, 0.5 g KCl, trace elements [172], 1.5 % agar, and 20 g fructose as a carbon source (pH 6.0). Protoplasts were transformed according to standard procedures [76]. *Escherichia coli* strain TOP10 (Invitrogen) was used as a host to propagate plasmids, and *E. coli* BL21(DE3)pLysS (Novagen) for expression of HacA.

DNA sequencing reactions were performed with the Bigdye sequencing kit (Applied Biosystems) and analysed with the ABI Prism 3100 Genetic Analyser (Applied Biosystems).

Table 1. PCR oligonucleotides used in this study

Oligo	Sequence
P1	5'-CCAGGCGTTCCTCATCGGTGAAAGCG
P2	5'-ACAAGCCAAGCACGACTCTGCCCAGAGG
P3	5'-CCTCCTCGGCGGTGCAGTCAACCTTCAC
P4	5'-GCTTTAGGTTCTGGCAGTGGCCGCACCA
P5	5'-CGGTGCTCGTCCGCATCAACCTTGCC
P6	5'-A<u>CCATG**G**</u>TGGAAGAAGCATTCTCTCCAG
P7	5'-**G<u>GGATCC</u>TAGTGGTGGTGGTGGTGGTG**TGCAGGATGTTGTGTCACGTC
P8	5'-CAGGTCAGTTCAGCGGATCCTGTCG-$(N)_{26}$-GAGGCGAATTCAGTGCAACTGCAGC
P9	5'-CAGGTCAGTTCAGCGGATCCTGTCG-$(N)_8$-C-$(N)_4$-TGNCCT-$(N)_5$-GAGGCGAATTCAGTGCAACTGCAGC
P10	5'-CAGGTCAGTTCAGCGGATCCTGTCG
P11	5'-GCTGCAGTTGCACTGAATTCGCCTC
P12	5'-CAGTAATGCAAGC**C**GGCCTCCATT
P13	5'-CTCCGCAGTAATG**AC**A**CG**TG**T**CCTCCATTCGCTG

Changes to the wild-type situation are shown in bold and introduced restriction sites are underlined.

5'-Random Amplification of cDNA Ends (5'RACE)

5'RACE was performed on 1 µg of total RNA using the SMART RACE cDNA Amplification Kit (Clonetech), according to the recommendations of the supplier. Primers P1, P2, P3, P4, and P5 (Table 1) were used to amplify *bipA, cypB, pdiA, prpA,* and *tigA* cDNA respectively.

Truncation of HacA

A fragment encoding the first 214 amino acids of HacA ($HacA_{(1-214)}$) was PCR amplified from $hacA^i$ cDNA using primers P6 and P7 ($hacA^i$ refers to the induced form of the *hacA* transcript, which lacks the 20 nt unconventional intron. $hacA^u$ refers to the unspliced, and therefore the uninduced form of the *hacA* mRNA). A C-terminal $(his)_6$-tag was included for purification purposes. After sequencing of the construct, a *NcoI-BamHI* fragment was ligated into plasmid pET3d (Novagen). *E. coli* BL21(DE3)pLysS (Novagen) was transformed with the resulting plasmid to express the truncated HacA protein. The protein was purified under native conditions by affinity chromatography using Ni-NTA agarose columns (Amersham Pharmacia Biotech). Protein concentrations were determined according to standard procedures [9] using bovine serum albumin (BSA) as a standard.

Electrophoretic Mobility Shift Assay (EMSA)

Oligonucleotide primers were annealed in 10 mM TRIS-HCl (pH 7.5) containing 100mM NaCl, by heating the mixtures to 95°C and cooling them down to room temperature. 1 pmol of annealed oligo was end-labelled with γ-^{32P}ATP using T4 Polynucleotide Kinase (New England Biolabs), and purified on Biogel P30 spin columns (BioRad). Binding reactions (20 µl) were performed in the following binding buffer: 25 mM TRIS-HCl (pH 8.0), 5 mM dithiothreitol, 100 mM KCl, 4 mM spermidine, 0.1 µg/µl poly[d(I-C)], 0.25% BSA, 5% glycerol, and included 1nM labelled probe, and variable amounts of purified HacA protein. Binding was allowed to proceed on ice for 15 min. The reactions were loaded on a 6% polyacrylamide gel and electrophoresed at 100 V in 0.25 x TBE (89 mM TRIS-HCl, 20 mM EDTA, 89 mM Boric acid, pH 8.3).

Selection for Oligonucleotides with High Affinity for HacA

Binding reactions were performed as described before, with random oligo P8 and the partly random oligonucleotide P9. The complementary DNA strand on both oligos was generated in a Klenow reaction together with oligo P11. To ensure stringent selection conditions, 10 nM of the *cypB* UPRE oligo was included in the binding reaction. After binding, the HacA-oligo complex was purified by affinity chromatography using Ni-NTA agarose mini columns. The columns were washed extensively with buffer containing 25 mM TRIS-HCl (pH 8.0), 5 mM dithiothreitol (DTT), 100 mM KCl, 4 mM spermidine, and 20 mM imidazole, and eluted with buffer containing 50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, and

250 mM imidazole. Bound oligonucleotides were amplified by PCR with primers P10 and P11. The amplified material was then used in a next cycle of selection and amplification. After 8 cycles, the selected oligonucleotides were cloned into pCR-Blunt II-TOPO (Invitrogen) and sequenced.

Estimation of the Relative Affinities

The apparent dissociation constant (K_d) between HacA and the various UPREs was calculated from DNA gel band shift reactions. At equilibrium the concentration of free protein when half of the probe is complexed, is equivalent to the apparent K_d [133]. Probe concentrations were 1 nM in all assays, and protein concentrations varied from 5 nM to 5μM. The apparent K_d is therefore given by the total concentration of protein at half the saturation of the probe minus 0.5 nM. Quantification was performed by phosphor-imager analysis (Packard; Instant Imager)

GUS Reporter Constructs

The GUS encoding gene (*uidA*) of *E. coli* was placed under control of a 280 bp fragment of the *cypB* promoter containing the putative UPRE. Mutations were introduced in the *cypB* UPRE by site directed mutagenesis using primers P12 (mutation 6T→C) and P13 (the binding site selection sequence, or BSS). A 3.9 kb *XbaI* fragment from plasmid pAB4.1, bearing a non-functional *pyrG** gene, was included in the pUPRE plasmid, which allowed for targeted integration of the construct at the *pyrG* locus of *A. niger* AB4.1 [169].

Southern Blot Analysis

Genomic DNA for Southern blot analysis was isolated by adding 800 μl buffer (100 mM TRIS-HCl (pH 8.0), 50 mM EDTA, 1 mM DTT, 35 mM SDS) to ground mycelium. The DNA was purified by phenol-chloroform extraction according to standard procedures [141]. The DNA pellet was resuspended in H_2O and incubated with DNase free RNase (Roche Molecular Biochemicals). For each digest 10 μg of genomic DNA was incubated 6 h at 37°C with 20U *Eco*RI, which cuts once in pUPRE. The reaction products were separated on a 0.8% TBE agarose gel. Blotting and hybridisation was done according to standard procedures [141].

GUS assay

A. niger AB4.1 reporter strains were grown overnight at 30°C in minimal medium supplemented with 0.2% w/v yeast extract. Aliquots of 1 g wet mycelium were transferred to 100 ml fresh medium and the cultures were incubated for 3 h before provoking ER-stress by the addition of either 20 μg/ml tunicamycin or 20 mM DTT. Samples from stressed and unstressed cultures were taken 1, 2, 4, 6, 8 and 10 h after the applying the stress agents, and flash frozen in liquid nitrogen. For total protein preparations, mycelium

was ground in reaction buffer (50 mM $NaPO_4$ pH 7.0, 10 mM DTT, 1 mM EDTA, 0.1% triton X-100) with 1mm glass beads using a Fast Prep FP120 (BIO 101 Savant). The total protein concentration in the cell extract was determined according to standard procedures [9]. The GUS activity was determined at 37°C in 1 ml reaction buffer containing 1 mM *p*-nitrophenyl-β-D-glucuronide. Reactions were stopped by the addition of 0.4 ml 1M Na_2CO_3, and the absorbance at 410 nm was determined. The activity in units was determined as the formation of *p*-nitrophenol in µmole per minute per µg of total protein. Readings from the untransformed parental strain *A. niger* AB4.1 were regarded as background and subtracted from the measurements. The experimental error in the calculated activity was derived by inserting the highest and lowest values of the 68% confidence interval of both the *p*-nitrophenol and the protein determination in the equation for the unit determination. Two activities are significantly different when the difference between them is larger than the sum of their standard errors.

Results

Identification of the ER Stress Response Element

Recently we showed that the 5'-flanking regions of the *A. niger* genes *bipA, cypB, pdiA prpA, tigA* and *hacA* all contained UPREs, since their promoter fragments were retarded in a gel band shift experiment with HacA [101]. Promoter deletion analysis narrowed down the location of the *cypB* UPRE to a 120 bp region flanking the 5'-end of the *cypB* gene (data not shown). Since the HacA homologues from *A. niger*, *A. nidulans* and *Aspergillus fumigatus* have an identical DNA binding domain, and therefore might recognise similar targets, a sequence comparison between the 5'-flanking regions of the *cypB* homologues from these three *Aspergilli* was made. This revealed a highly conserved box, present within the 120 bp 5'-flanking region of the *cypB* gene (Fig. 1A). Purified HacA protein was able to bind *in vitro* to a 34 bp DNA oligonucleotide containing the putative *cypB* UPRE, deduced from sequence alignments (Fig. 2A). This binding proved to be specific because even a 500-fold excess of an unlabelled probe containing a point mutant of the *cypB* UPRE did not compete significantly in the binding assay.

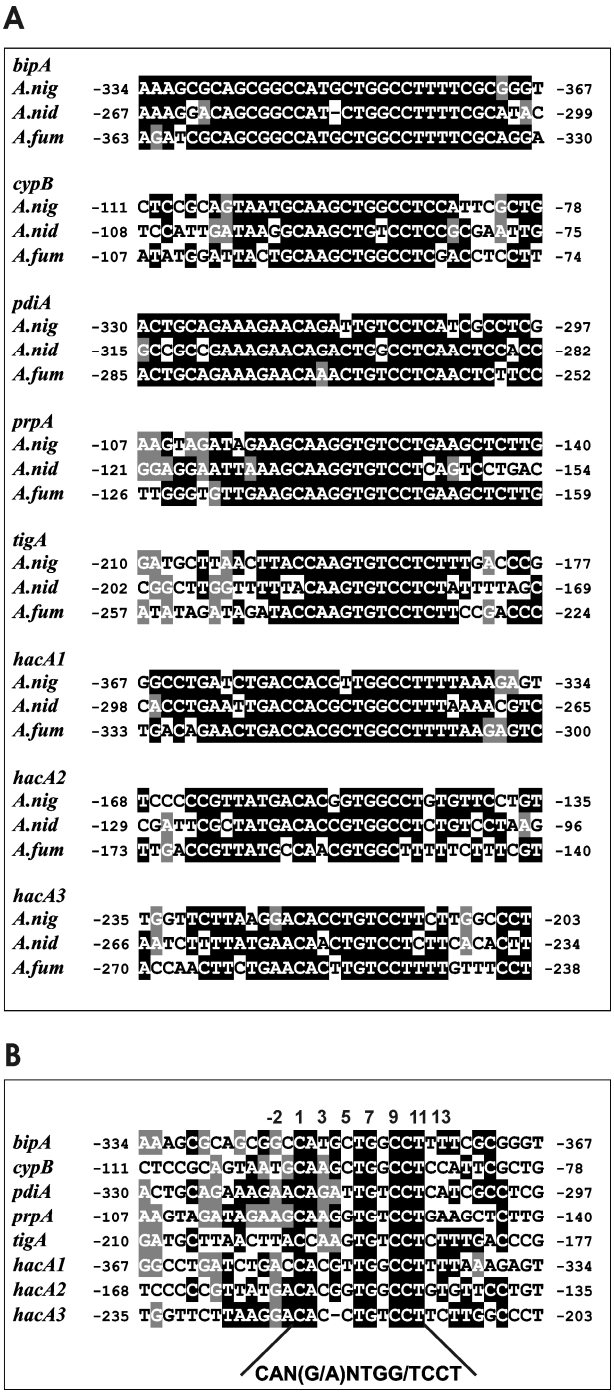


Figure 1. Identification of putative Unfolded Protein Response Elements (UPREs) from *Aspergilli*. **(A)** Alignment of conserved promoter regions that shared sequence similarity with the *A. niger cypB* UPRE. Alignments are triple wise among the homologues of three *Aspergilli*; *A. niger (A.nig)*, *A. nidulans (A.nid)*, and *A. fumigatus (A.fum).* Numbers beside the sequences indicate the position relative to the ATG of the downstream located gene. The 11 bp core sequences of the UPREs are underlined. **(B)** Alignment of the UPREs identified in the promoters of the *A. niger* genes *bipA, cypB, pdiA, prpA, tigA*, and *hacA*. The bases of the conserved core are numbered above the alignment. The 11 bp consensus sequence is indicated under the alignment

To identify putative UPREs in the promoters of *bipA, pdiA, prpA, tigA* and *hacA,* a similar approach was followed. Blast searches were performed to retrieve non-annotated homologues from the *A. fumigatus* and *A. nidulans* genome databases (http://www.tigr.org/tdb/e2k1/afu1/ and http://www.broad.mit.edu/annotation/fungi/aspergillus/index.html).

Each set of homologous promoters from the three *Aspergilli* was analysed for the presence of conserved blocks using the program MACAW [145]. Conserved boxes with similarities to the *cypB* UPRE were found in all promoters (Fig. 1A). Outside these boxes, the sequence identity is much lower, which is most obvious from the alignments of the *cypB* and *tigA* promoter regions. EMSA showed that all sequences identified as potential UPREs were bound by HacA (Fig. 2B). Interestingly, two putative UPREs (indicated as *hacA1* and *hacA2*) were identified in the promoters of the *hacA* homologues (Fig. 1A). The fact that both UPREs are well conserved among the three *Aspergilli*, and the observation that both *A. niger hacA* UPREs are bound by HacA *in vitro* (Fig. 2B), could indicate an *in vivo* function for both boxes. (The third UPRE in the *hacA* promoter, *hacA3,* was discovered later, and will be discussed separate, later on in the results section). Figure1B shows the alignment of the seven *A. niger* UPREs identified in this study. The consensus sequence, CAN(G/A**)**NTG**T/GCCT**, uncovered by this alignment shows similarities with the *S. cerevisiae* UPRE. The *S. cerevisiae KAR2* UPRE, CAG(C)GTG, is a partly palindromic sequence separated by a spacer of one nucleotide [97]. The palindromic residues CA and TG (underlined) are common for a number of *S. cerevisiae* UPREs, and the spacer nucleotide C (between brackets) is conserved in almost all *S. cerevisiae* UPREs. From the alignment of the UPREs found in the promoters of several *A. niger* genes it is clear that the *A. niger* UPRE also has partial symmetry around a spacer nucleotide. However, in contrast to *S. cerevisiae*, in the *A. niger* UPREs the spacer is either G or A. Furthermore, the partly palindromic box in *A. niger* is followed by the highly conserved sequence **T/GCCT** (in bold), which is not found in *S. cerevisiae*.

The alignments shown in Figure 1 revealed that the UPREs are remarkably well conserved among homologous genes across the species boundary, whereas more variation occurs among the UPREs of the different UPR regulated genes within a single *Aspergillus* species. For example, the 34 bp regions containing the UPREs of the *pdiA*, *prpA*, and *tigA* homologues (Fig. 1A) share 68%, 50%, and 38% identical residues, respectively. However, among the UPREs of these three protein disulphide isomerases of *A. niger* (Fig.1B), there are only 29% identical residues.

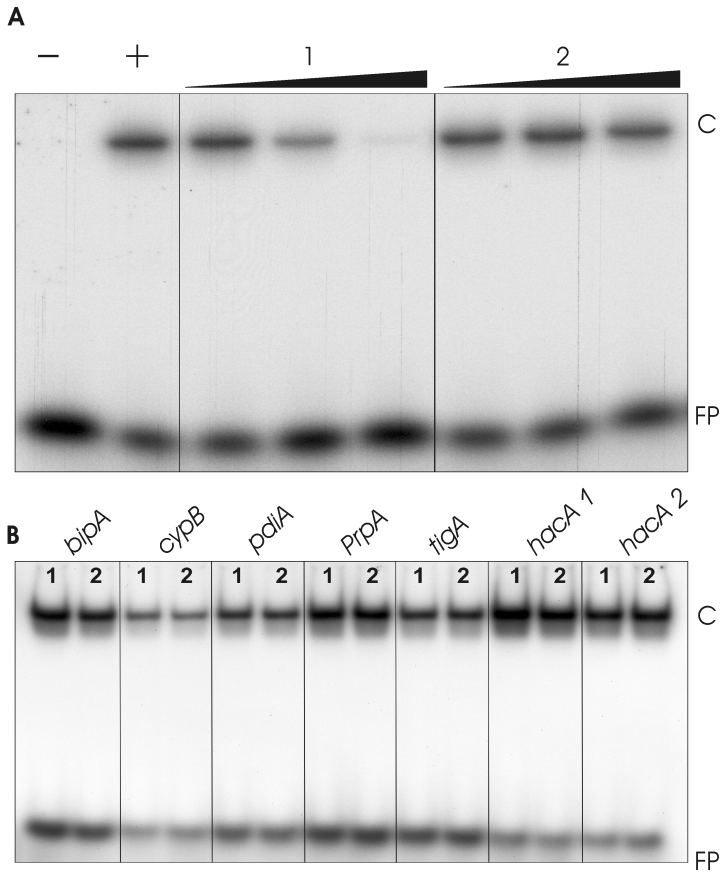


Figure 2. Electrophoretic mobility shift assay (EMSA) showing the binding of the *A. niger* HacA to 34 bp oligonucleotides encompassing the identified UPREs. The binding reactions were performed with 200 nM HacA protein and 1 nM ^{32}P-labelled probe. **(A)** Specific binding of HacA to the *cypB* UPRE. (-) without HACA. (+) with HACA. (1) Competition with 5, 50 and 500 nM unlabelled *cypB* UPRE, respectively . (2) Competition with 5, 50 and 500 nM unlabelled mutant *cypB* UPRE. The mutant *cypB* UPRE contained the mutation 6T→C (cf Table 2). **(B)** Binding of HacA to the *bipA, cypB, pdiA, prpA, tigA, hacA1, and hacA2* UPREs. (1) Without competition. (2) Competition with 500 nM unlabelled mutant *cypB* UPRE. C-HacA-DNA complex, FP-Free probe.

Mutation Analysis of the *cypB* UPRE

To evaluate the contribution of each nucleotide of the consensus sequence to DNA binding, we assessed the effect of 14 mutations at different positions in the *cypB* UPRE (Table 2). The mutations 1C→A and 2A→C, both located in the left half-site of the imperfect palindrome, resulted in a 2.5-fold decrease and a 2-fold increase in binding, respectively. In contrast, mutation of the palindromic residues in the right half-site of the imperfect palindrome completely abolished binding (mutations 6T→C, 6T→G and 7G→T). As stated before, most *S. cerevisiae* UPREs have a spacer nucleotide, usually a C, interrupting the palindrome. At the corresponding position in the UPRE of *Aspergillus,* usually a G or an A is present. Although no T was found at the position of the spacer nucleotide, the mutation

4G→T resulted in a substantial increase in binding of the *cypB* UPRE. Position 8 in the *Aspergillus* UPRE is occupied by either a T or a G (Fig. 1B), and mutation 8G→T in the *cypB* UPRE only had a minor effect on binding. However, binding was completely abolished when this residue was mutated to a C. Changes in the last three conserved residues (CCT) had dramatic effects. Binding is lost by the mutation 9C→A, and weakened substantially by the mutations 10C→A, 11T→C, and 11T→G.

Table 2. Summary of the gel shift analysis of double-stranded DNA oligonucleotides representing the cypB UPRE and containing mutations in the core binding sequence.

Probe	Sequence	% of binding relative to wild-type probe
wt	5'-CAGTAATG AAAGCTGG CCTCCATT	100
1A	5'-CAGTAATG **A**AAGCTGG CCTCCATT	40
2C	5'-CAGTAATG C**C**AGCTGG CCTCCATT	209
3C	5'-CAGTAATG CA**C**GCTGG CCTCCATT	6
4T	5'-CAGTAATG CAA**T**CTGG CCTCCATT	169
5G	5'-CAGTAATG CAAG**G**TGG CCTCCATT	155
6C	5'-CAGTAATG CAAGG**C**GG CCTCCATT	0
6G	5'-CAGTAATG CAAGG**G**GG CCTCCATT	0*
7T	5'-CAGTAATG CAAGCT**T**G CCTCCATT	0
8T	5'-CAGTAATG CAAGCTG**T** CCTCCATT	81
8C	5'-CAGTAATG CAAGCTG**C** CCTCCATT	0*
9A	5'-CAGTAATG CAAGCTGG **A**CTCCATT	0
10A	5'-CAGTAATG CAAGCTGG C**A**TCCATT	39
11C	5'-CAGTAATG CAAGCTGG CC**C**CCATT	21
11G	5'-CAGTAATG CAAGCTGG CC**G**CCATT	15*

Mutations are indicated in **bold**. The binding relative to the wild-type was determined based on apparent *Kd* values, except for those marked with and *asterisk*, which were determined from single DNA bandshift experiments.

Selection for High Affinity Binding Sites

Binding site selection is a rapid method to isolate oligonucleotides with high affinity for DNA binding proteins [128]. To determine the optimal binding site for HacA, a binding site selection strategy was employed. Binding reactions were performed with a 74 bp double-stranded oligonucleotide including 25 bp flanks for PCR amplification. The central 24 bp were partly randomised. Bases which were conserved in all *A. niger* UPREs (1C, 6T, 7G, 9C, 10C and 11T) were kept at their original positions (cf Fig. 1B and P9 in Table 2). The experiment was performed in parallel with a double stranded oligonucleotide for which the central part was completely randomised. To ensure stringent selection conditions, the binding reaction included 10 nM of the 34 bp *cypB* UPRE as competitor DNA. A sequence alignment of the oligonucleotides selected after eight rounds of the selection procedure is shown in Figure 3.

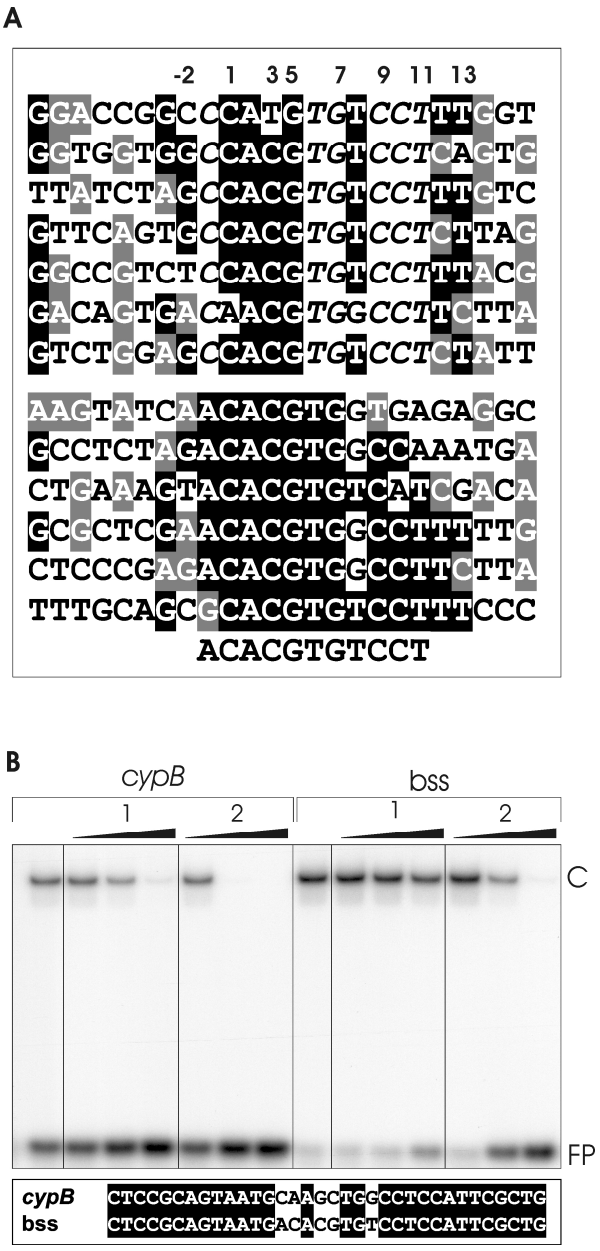


Figure 3. Selection of binding sites from random oligonucleotides with HACA-$(his)_6$. **(A)** Upper panel; Sequences selected with the partly random oligonucleotide P9. The fixed nucleotides in this probe, representing the conserved bases in the *A. niger* UPREs (cf Fig. 1B), are depicted in italics and are not accounted for in the alignment. Lower panel; Sequences selected with random oligo P8. Below the alignment are indicated the nucleotides most frequently present at each position (more than 60%). The bases included in the conserved core are numbered analogous to Figure 1. Note that residue nr. 4, the spacer nucleotide is absent. **(B)** EMSA showing the binding properties of the binding site selection box (BSS) compared to the *cypB* UPRE. On the left, binding of HacA to ^{32}P-labelled *cypB* UPRE followed by competition with (1) 5, 50 and 500 nM unlabelled mutant *cypB* UPRE, and (2) 5, 50 and 500 nM unlabelled BSS. On the right, binding of HacA to ^{32}P-labelled BSS followed by competition with unlabelled *cypB*, and unlabelled BSS. The *cypB* and BSS oligonucleotides

Interestingly, a perfect palindromic sequence, CACGTG, was present in nearly all cases. Strikingly, a spacer nucleotide between the two palindromic half-sites was absent in all of the selected oligos. Besides the lack of a spacer nucleotide and its perfect palindromic nature, the selected box (further referred to as BSS for binding site selection sequence) resembles the UPREs found in *A. niger*. The palindromic nucleotides CA and TG at positions 1-2 and 6-7 respectively, were found both in the UPREs and in the BSSs. Position 8 in the BSS are in all cases occupied by either a G or T nucleotide, similar to position 8 in the UPREs. Furthermore, the conserved stretch CCT at position 9-11 in the *A. niger* UPREs, was also frequently found in the BSSs. At position –1, directly preceding the palindromic core, an A nucleotide was present most frequently. This could be palindromic with the nucleotide at position 8, which is most frequently a T, thereby extending the palindromic core to eight bases.

From these data, it also appears that HacA prefers pyrimidines at position 12 and 13 and purines at position –2 and –3. The purines G and A were present at positions –2 and –3, respectively in nine and ten of the thirteen BSSs. The pyrimidines C and T were present at positions 12 and 13, in eleven and nine cases respectively. Similarly, G and A nucleotides were more frequently present at positions –3 and –2, and C and T nucleotides at positions 12 and 13 of the UPREs (Fig. 1).

Based on the above findings, five mutations were introduced in the *cypB* UPRE to change its partly palindromic core towards the BSS consensus sequence (Fig. 3B, bp number 1-5). The binding properties of this BSS were compared with the wild type *cypB* UPRE by EMSA analysis. Non-labelled BSS out-competed the *cypB* UPRE for binding by HacA. Conversely, an excess of the cold *cypB* UPRE was unable to significantly decrease the intensity of the HacA-BSS complex. This indicates a higher binding affinity for the BSS site compared to the *cypB* UPRE.

The *hacA* UPRE 3

The UPRE isolated by binding site selection lacked a spacer nucleotide, and was bound stronger by HacA than the *cypB* UPRE (Fig. 3B). Promoter sequences of *A. niger*, *A. nidulans* and *A. fumigatus* were therefore analysed for the presence of BSS-like UPREs. Surprisingly, such an UPRE was found in the *hacA* promoter itself (Fig. 1). Like the other UPREs, its sequence and position in the promoter is highly conserved among the 3 *Aspergilli*. Its consensus sequence is CACCTG**TCCT,** which is nearly identical to the BSS, and like the BSS the core sequence is preceded by several purine residues, and followed by several pyrimidines

Binding Affinities of the UPREs

Figure 4 shows a typical pattern of an EMSA obtained with increasing concentrations of HacA protein and 1 nM *cypB* UPRE. Below the EMSA is a graph where the saturation (R) is plotted against the concentration of HacA. Since HacA was found to bind DNA as a homodimer (see later this paper), the saturation (R) of the UPRE was defined as:

$$R=[HACA_2 \cdot UPRE]/([HACA_2 \cdot UPRE]+[UPRE])$$

For determination of K_d, the data was fitted to the following equation:

$$R=0.5[HACA]\ R_{max}\ /((K_d+0.5)\ 0.5[HACA]).$$

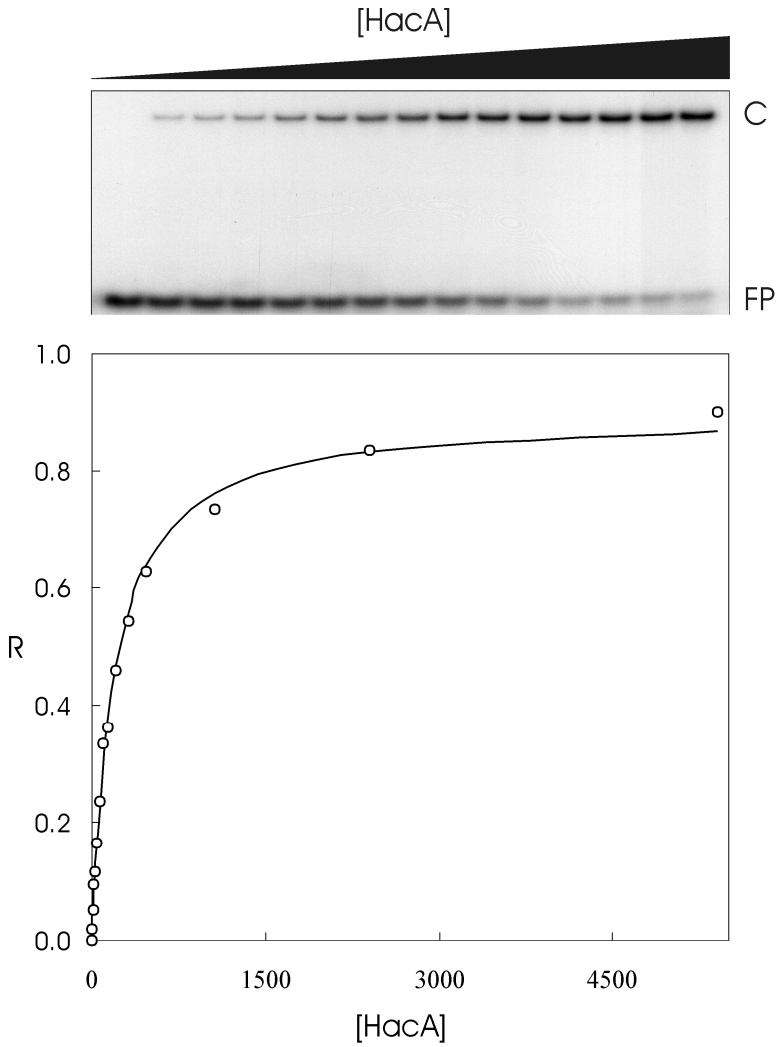


Figure 4. Saturation of the 34 bp *cypB* UPRE (R) versus the concentration of HacA. The K_d was calculated by fitting the data to equation above using SlideWrite Plus version 3.00. K_d values for the different *A. niger* UPREs are listed in Table 3.

The dissociation constants for the $[HACA]_2$-DNA complexes for the different UPREs and the BSS are listed in Table 3. The K_d values for the UPREs of *bipA, pdiA, prpA* and *tigA* are rather similar at approximately 30 nM. This is comparable to the apparent K_d of 35 nM reported for *S. cerevisiae* transcription factor GCN4 [35]. The K_d value for the *cypB* UPRE was found to be much higher at 96 nM, which was also evident from Figure 2B. In contrast, the three *hacA* UPREs have K_d values much lower than those of the UPREs present in the promoters of the foldases. Whether this reflects a physiological role remains to be investigated. Interestingly, the *hacA3* UPRE and the BSS, which both lack the spacer nucleotide, have the lowest K_d values of the UPREs that were analysed here. Compared to the *cypB* UPRE the artificial BSS UPRE has a drastically increased binding affinity. Five mutations were introduced in the *cypB* UPRE to obtain the BSS, but with a K_d value of 1.9 nM, its binding affinity is approximately 50 times higher.

Table 3. Dissociation constants of the *A. niger* UPRE-HacA complexes studied here

UPRE	***Kd* in nM**
bipA	28.9 ± 1.3
cypB	96.0 ± 5.0
pdiA	31.6 ± 2.4
prpA	32.5 ± 1.5
tigA	34.8 ± 1.5
hacA 1	10.9 ± 0.5
hacA 2	14.2 ± 0.7
hacA 3	5.7 ± 0.5
BSS	1.9 ± 0.1

In vivo Activity of the *cypB* and the BSS UPR Element

The *in vivo* activity of the *cypB* UPRE was assessed with reporter constructs having the *uidA* gene under control of a 280 bp promoter fragment of the *A. niger cypB* gene (pUPRE-wt). Changes were made in the UPRE of the *cypB* promoter, which resulted in three additional constructs: 'pUPRE-mut', containing mutation 6T→C which led to a complete loss of HacA binding in DNA gel shifts (cf Table 2), 'pUPRE-bss', containing the binding site selection box, and 'pUPRE-bss5x', containing the binding site selection box repeated five times (head to tail) in the *cypB* promoter. *A. niger* AB4.1 transformants having the reporter construct as a single copy integrated at the *pyrG* locus were selected by Southern analysis (data not shown). The reporter strains were grown in liquid media, ER stress was provoked by the addition of tunicamycin and DTT, and the development of intracellular β-glucuronidase activity was followed for 10 h (Fig. 5).

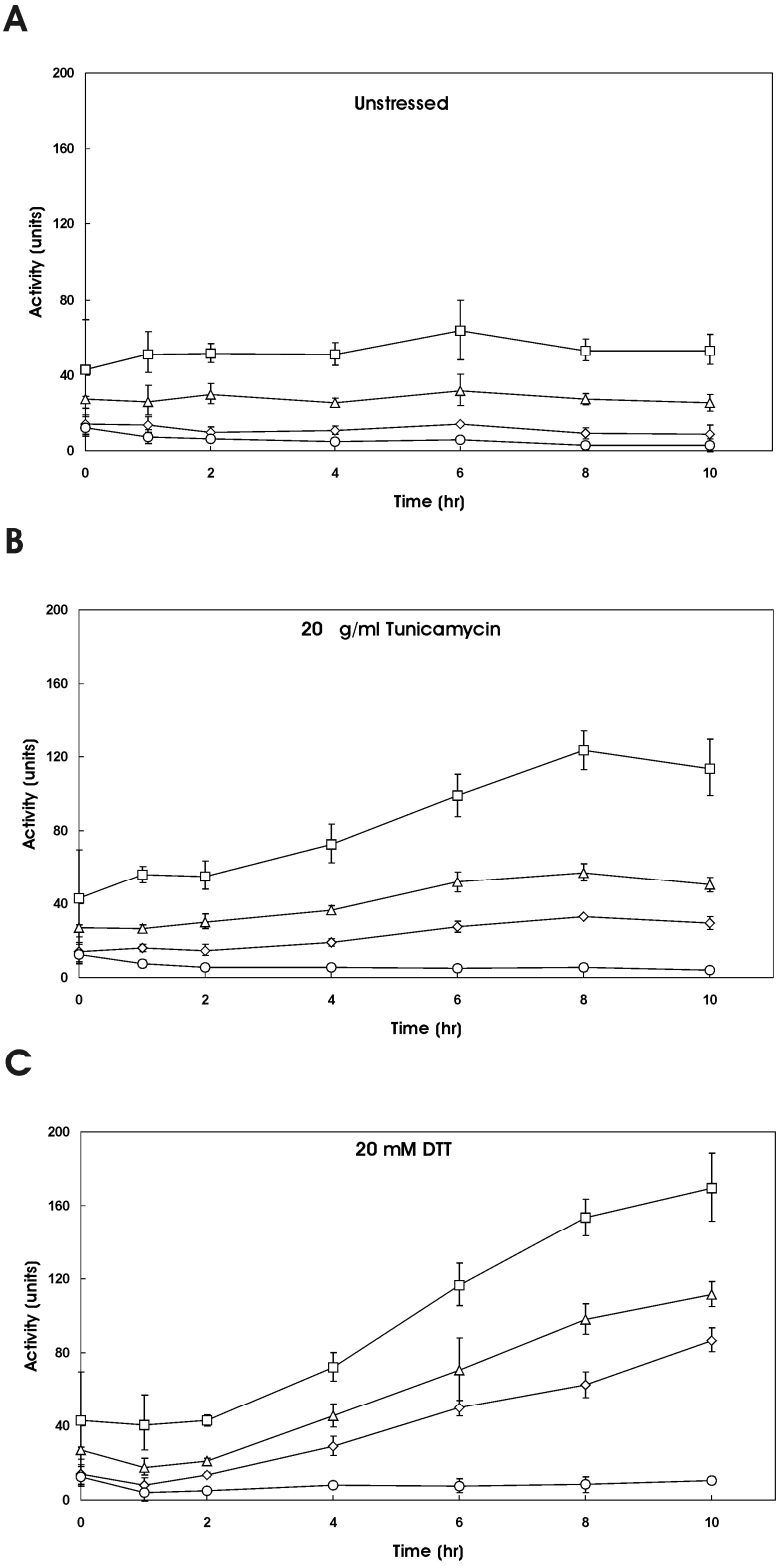


Figure 5. *in vivo* analysis of the UPRE in *A. niger*. Reporter strains contained a single copy of the construct pUPRE integrated at the *pyrG* locus. Intracellular β-glucuronidase activity and standard errors were determined as mentioned in the Experimental procedures. ○-*A. niger* AB4.1::pUPRE-mut, ◇- *A. niger* AB4.1::pUPRE-wt, △-*A. niger* AB4.1::pUPRE-bss, □-*A. niger* AB4.1::pUPRE-bss5x. **(A)** Development of β-glucuronidase activity in unstressed cultures, **(B)** in cultures treated with 20 µg/ml tunicamycin, and **(C)** in cultures treated with 20 mM DTT.

Compared to *A. niger* AB4.1::pUPRE-wt, the basal expression level was elevated considerably in the reporter strains carrying either pUPRE-bss or pUPRE-bss5x (Fig. 5A). *A. niger* AB4.1::pUPRE-mut on the other hand, exhibited a slightly lower basal activity compared to the pUPRE-wt. Upon ER stress, triggered by tunicamycin or DTT, the β-glucuronidase activity increased significantly in strains carrying pUPRE-wt, pUPRE-bss, and pUPRE-bss5x. The effect was most pronounced for pUPRE-bss5x, followed by pUPRE-bss and pUPRE-wt. No response was observed for pUPRE-mut, neither on tunicamycin nor on DTT. pUPRE-bss5x showed the largest increase in activity, both with tunicamycin and DTT. This indicates that the effect of at least some of the additional copies of the UPRE in this promoter is cumulative. In all cases, DTT gave a stronger response than tunicamycin. Tunicamycin specifically provokes ER-stress by inhibiting N-glycosylation of newly synthesised proteins, thereby blocking their folding and transit through the ER [122]. DTT acts more generally by disrupting the oxidising environment of the ER and thereby preventing the formation of disulphide bonds. Taken together, these results demonstrate that the HacA DNA-binding sites established *in vitro,* are of physiological relevance.

HacA Binds DNA as a Dimer

The consensus UPRE of *A. niger* was found to be a partly palindromic sequence, and a perfect palindromic box was isolated by a binding site selection approach. Furthermore, HacA exhibits features typical for the family of basic leucine-zipper type transcription factors (bZIP). Taken together this suggests that HacA binds to DNA as a homodimer. To assess whether HacA indeed binds to the UPRE as a dimer, a truncated form of HacA was expressed in *E. coli.* A stop codon was introduced at the start of the unconventional intron of *hacA*, which resulted in a HacA variant lacking 128 amino acids at the C-terminus. This truncated HacA was used in a DNA binding experiment together with the full length HACA, and the 34-bp ^{32}P-labelled *cypB* UPRE (Fig. 6).

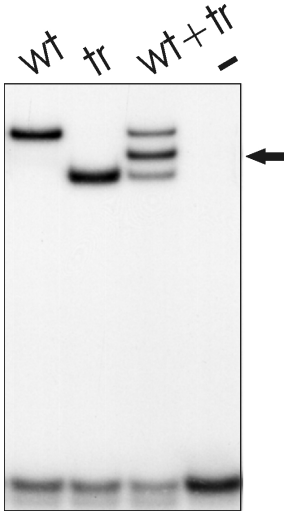


Figure 6. EMSA showing that HacA binds *in vitro* to DNA as a homodimer. (wt), the full length HACA-$(His)_6$. (tr), the truncated $HACA_{1\text{-}214}$-$(His)_6$. The position of the dimer, formed between the full length and the truncated HacA is indicated by an arrow. 1 nM ^{32}P-labelled *cypB* UPRE was used as a probe.

The complexes formed between the *cypB* UPRE and either the full length or truncated HacA proteins, had different migration rates but possessed approximately the same intensity, which indicates similar binding properties. When both HacA proteins were added to the same binding reaction an additional complex of an intermediate size was formed, as judged by the EMSA analysis (Fig. 6). This complex corresponds to a dimer formed between one full length and one truncated molecule of HacA, confirming that it functions as a dimer. Since the distribution of the three complexes is also as expected, approximately 25% full length HacA dimer, 50% full length HACA/truncated HacA dimer, and 25% truncated HacA dimer, this indicates that the C-terminus of HacA does not significantly affect the DNA binding properties of the protein *in vitro.*

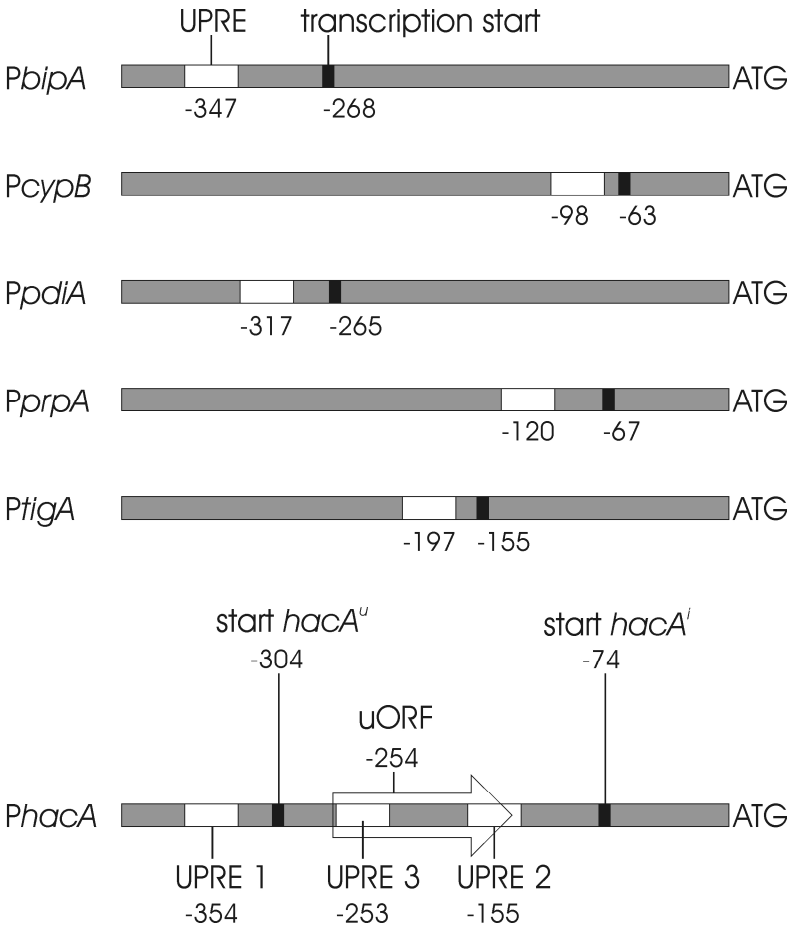


Figure 7. Schematic representation of the position of the *A. niger* UPREs in relation to the putative transcription start site, and the ATG codon of the corresponding gene. Numbers indicate the position of the transcription start, and the position of the conserved C at position 1 of the UPRE (cf Fig. 1) relative to the ATG codon. The transcription starts are indicated by black boxes, and the UPREs by open boxes. *hacA*^{i} refers to the induced form of the *hacA* transcript, *hacA*^{u} to the uninduced form. The open arrow indicates the small open reading frame, present in the 5'UTR of *hacA*^{u}.

Determination of the Distance Between UPR-elements and Transcription Start Sites

The distance between the UPREs in the 5'-flanking regions and the corresponding ATG codons, in different UPR-controlled genes, varies between 334 nt (*bipA*) and 78 nt (*cypB*). To determine whether there is less variation between the positions of these HacA binding sites in relation to the transcription start site, 5'RACE experiments were performed for five target genes. Figure 7 shows schematically the position of the UPRE in relation to both the transcription start site and the ATG-codon in a number of HacA regulated promoters.

In all cases, the UPRE is in close vicinity to the transcription start site. The distance between the conserved C at position 1 of the UPREs (cf Fig. 1B) and the transcription start site varies between 79 bp for *bipA* and 25 bp for *cypB*. Previously we determined the start sites of the *hacA*^{u} and *hacA*^{i} transcripts [101]. It turned out that these start sites are preceded by UPRE1 and UPRE2 respectively, whereas UPRE3 is situated in between the two other UPREs, and overlaps with the start of the uORF (Fig. 7).

Discussion

Recently we cloned and characterised the *A. niger* transcription factor HacA, involved in signalling the presence of ER stress to the nucleus. Over-expression of *hacA*^{i} led to the up-regulation of UPR target genes, and DNA band shift analysis with the promoter fragments of those genes confirmed the direct interaction of HacA with those promoters [101]. Here we identified the *cis*-acting elements, present in the promoters of these UPR responsive genes of *A. niger*. The *A. niger* UPRE maps closely to the corresponding transcription start point (Fig. 7), and matches the following consensus sequence stretch: CAN(G/A)NTG**T/GCCT.** A partly palindromic sequence (underlined) around a single spacer nucleotide (between brackets), followed by a highly conserved stretch of four bases (in bold). The palindromic core of the consensus *A. niger* UPRE shows similarities to the *KAR2* UPRE of *S. cerevisiae*, which is: CAG(C)GTG [97,99]. Both UPREs contain an E-box like palindrome (CANNTG) around a single spacer nucleotide [44]. The similarity between the *A. niger* and *S. cerevisiae* UPREs is also reflected by several other findings. Firstly, although the overall identity between the *A. niger* HacA and *S. cerevisiae* Hac1p is only 25%, they share 77% identical residues in their DNA binding domains [101]. Secondly, the *A. niger* HacA was found to bind an oligonucleotide representing the *S. cerevisiae KAR2* UPRE in a DNA gel shift assay (data not shown). Furthermore, the *A. niger* HacA shares 27% overall identity and 92% identical residues in the DNA binding domain with the *Trichoderma reesei* Hac1 protein. The *T. reesei hac1* gene could partly complement the *S. cerevisiae HAC1* disruption [140], and over-expression of the *T. reesei hac1* in *S. cerevisiae* resulted in accumulation of the *KAR2* transcript, which indicates recognition of the *S. cerevisiae KAR2* UPRE by the *T. reesei* Hac1 protein [164].

Although the *A. niger* and the *S. cerevisiae* UPRE share some common features, there are three major differences; the nature of the spacer nucleotide, the highly conserved stretch following the palindromic core of the *Aspergillus* UPRE, and the dispensability of the spacer nucleotide.

The first notable difference between the *A. niger* and *S. cerevisiae* UPRE is the nature of the spacer nucleotide, which in the UPRE of *S. cerevisiae* most frequently is a C nucleotide, and Hac1p prefers a specific spacer nucleotide in the order C>G>A>T [95]. In the *A. niger* UPREs, however, a G or an A was present as spacer nucleotide, given the situation in Figure 1 in which the majority of the aligned UPREs are shown in their natural orientation. Moreover, the mutation 4G→T of the spacer nucleotide in the *cypB* UPRE led to significantly increased binding (Table 2), this in contrast to the order of preference reported for the *S. cerevisiae* Hac1p.

The second important difference is the presence of a highly conserved stretch (T/GCCT) adjacent to the palindromic core of the *Aspergillus* UPREs, which appears to be absent in *S. cerevisiae*. This stretch was found in all *A. niger* UPREs analysed, and with one exception (GCTT in the *A. fumigatus tigA* UPRE) in the homologues from *A. nidulans* and *A. fumigatus* as well (Fig 1). The importance of these nucleotides for binding of HacA to the UPRE, was shown by mutational analysis of the *cypB* UPRE, demonstrating that mutations in any of the last three residues of the stretch (CCT) affected binding significantly (Table 2.). Furthermore, binding site selection also led to a box having this stretch, following the palindromic core (Fig. 3).

The third difference is the dispensability of the spacer nucleotide in the *Aspergillus* UPRE. Deletion of the spacer in the *S. cerevisiae KAR2* UPRE reduced the UPR response to 13% of the wild type level, and introduction of two or more C bases between the palindromic half-sites abolished Hac1p-mediated transcriptional activation almost completely [95]. In contrast to the situation reported for the *S. cerevisiae* Hac1p, the *A niger* HacA prefers an UPRE without a spacer nucleotide. Although most of the UPREs identified in the promoter regions of the HacA-controlled genes in *A. niger* actually contain a spacer nucleotide, we demonstrated that despite the lack of it, the binding site selection box operates as a physiologically relevant *cis*-acting element when introduced in *A. niger* (Fig. 5). This suggests that an UPRE without a spacer nucleotide could be of biological relevance, and several other findings strengthen this hypothesis. Firstly, the *hacA* promoter itself contains such an UPRE (Fig. 1). The transcription of *hacA* is auto-regulated [101], and transcriptome analysis of the UPR in *A. nidulans* puts *hacA* among the highest up-regulated genes [154], which is in accordance with our finding of the *hacA* UPRE3 having the lowest K_d value (Table 3). It should be noted, however, that the *hacA* promoter contains 3 UPREs, all having lower K_d values than the other UPREs studied here. Therefore, mutation analysis should be carried out, to determine the contribution of the different UPREs, and especially the UPRE3, to the transcriptional regulation of *hacA*. Secondly, the *A. nidulans bipA* UPRE

lacks the spacer nucleotide (cf Fig. 1). Except for the missing spacer nucleotide, its sequence (from –7 to +17) is identical to its homologues from *A. niger* and *A. fumigatus*. The binding affinity of the *A. nidulans bipA* UPRE (K_d =39 ± 9 nM) is in the same range of the binding affinities of the *A. niger* UPREs (Table 3), and only slightly lower compared to the *A. niger bipA* UPRE. This indicates that the spacer nucleotide is dispensable, at least in some *Aspergillus* UPREs. Finally, a closer inspection of the *A. niger tigA* UPRE revealed that it also might lack a spacer nucleotide. Its core sequence can be presented with a spacer nucleotide as CCA(A)GTG**TCCT**, but more likely without a spacer nucleotide as CAAGTG**TCCT**. Furthermore, the fact that the *A. nidulans tigA* UPRE only in the latter case possesses palindromic residues (Fig. 1A) suggests that the situation without a spacer nucleotide is more likely. However, mutational analysis should be carried out to unambiguously prove the lack of a spacer nucleotide in the *tigA* UPRE. Combined, these findings strongly argue for the dispensability of the spacer nucleotide, and the biological relevance of such an UPRE.

The UPREs from *A. niger, A. nidulans* and *A. fumigatus*, are remarkably well conserved among homologous genes across the species, while more sequence variation is present among the UPREs of different genes from *A. niger* (Fig. 1). The dissociation constants (K_d) between several of the *A. niger* UPREs also vary significantly (Table 3). Taken together this might indicate that during the course of evolution, the need for differential up-regulation of the different UPR-target proteins upon ER stress, has kept the UPREs of homologous genes rather similar. *In vivo* analyses showed that binding-enhancing changes in the *cypB* UPRE elevated the expression of GUS, also under non-stress conditions. In addition, the basal expression level was decreased by a mutation that inactivated the UPRE (Fig. 5). This suggests that under conditions not considered as stress, the transcription of UPR target genes is still, to some extent, regulated by the UPRE, implying the presence of a small amount of active HacA protein. An alternative hypothesis implies involvement of another, yet unknown, transcription factor that would also bind to the UPRE in *A. niger*.

The relative binding affinity of the BSS UPRE is approximately 50-fold higher than the *cypB* UPRE (Table 3), and this difference seems not to be reflected by the reporter assays (Fig. 5B and C). However, a plausible explanation could be that in cells stressed with either DTT or tunicamycin, the ratio of HacA molecules over UPRE targets might be that high, that the system is operating with an HacA concentration above the K_d values of the different UPREs. In such a situation, UPREs with different binding affinities would be occupied by HacA for comparable periods of time (cf Fig. 4), and therefore result in only minor differences in expression levels (Fig. 5). DTT and tunicamycin are agents known to provoke a strong UPR, and are therefore used to study ER stress. However, a biological stress response might not always lead to such a strong UPR, and small increases in the

HacA concentration might therefore result in differential expression of UPR target genes, depending on the K_d values of their UPREs.

HacA belongs to the basic leucine-zipper family of proteins. bZIP proteins normally have a DNA binding domain rich in basic amino acids followed by a coiled-coil dimerisation motif of a leucine-zipper type, and they are known to bind DNA either as homo- or hetero-dimers [78]. The partly palindromic core of the UPREs identified, points towards a symmetrically organised binding factor (Fig. 1), and DNA binding experiments with HacA polypeptides of different lengths confirmed that, *in vitro* at least, the protein binds to its cognate DNA targets as a homodimer (Fig. 6). However, recent work in *S. cerevisiae* revealed the importance of Gcn4p for the UPR, and Hac1p is thought to act in conjunction with Gcn4p, and stimulate transcription by binding to promoter regions as a Hac1p/Gcn4p heterodimer [124]. Whether the *A. niger* Gcn4p homologue CpcA [177] would fulfil a similar role in the UPR remains to be investigated. However, if HacA would operate as a heterodimeric complex with CpcA, it seems likely that HacA would recognise and bind the right half-site of the UPRE. This is the most conserved part of the UPRE, and mutations in the right half-site of the *cypB* UPRE had more dramatic effects on the binding of HacA than mutations in the left half-site (Table 2). Furthermore, the BSS has a right half-site which is identical to the right half-site of the consensus UPRE (cf Fig 1 and 3). The BSS has *cis*-activating properties *in vivo,* but was selected using an *in vitro* procedure with purified (homodimeric) HacA protein, which indicates the importance of the right half-site for HacA binding.

The central activation mechanism of the UPR is conserved among *S. cerevisiae*, mammalian cells and filamentous fungi, and comprises the splicing of an unconventional intron from the transcripts of *HAC1*, *XBP1* and *hacA* respectively. In contrast to its homologues in *S. cerevisiae* and mammalian cells, the *A. niger hacA* mRNA is shortened at its 5'UTR by 230 nt upon ER stress. This removes an uORF from the *hacA* mRNA, which could serve as an additional regulation step in the activation of the UPR in *Aspergillus*. Small uORFs can act as regulatory elements that control the translation initiation of the downstream located gene [171], and removal of the uORF could relieve the *hacA* transcript from translational attenuation, allowing for efficient translation during ER stress. The mechanism behind this 5'-truncation of the *hacA* transcript is, however, still unknown, but it either could be the result of a nuclease assisted processing event, or the result of a new transcription start site upon stress. However, the presence of three UPREs in the *hacA* promoter, and their arrangement in relation to the start sites of $hacA^u$ and $hacA^i$, argue for the latter case (Fig. 7). This would imply an ingenious auto-regulation mechanism, and we are currently studying the possible role of the different UPREs and the uORF on the regulation of *hacA* in *A. niger,* by mutational analyses.

Optimisation of the secretion machinery for heterologous protein expression requires the identification and relieve of bottlenecks or problems in the production process. A good understanding of the secretion pathway is therefore necessary. Knowledge

about the UPRE presented here could be used in the data mining of genome sequences for putative target genes. In combination with genome wide expression analyses of strains over-expressing HACA, this could lead to the identification of new HacA targets. Some of these might play important roles in the secretion machinery of filamentous fungi, and could be targets for future optimisation.

Acknowledgement

This project was part of the Eurofung program and supported by EC Grant No. QLRK3-00729.

Chapter 4

A HacA-Dependent Transcriptional Switch Releases *hacA* mRNA from a Translational Block upon Endoplasmic Reticulum Stress

Harm J. Mulder, Igor Nikolaev

Published in Eukaryotic Cell. 2009, 8(4): 665-675

Abstract

Activation of the unfolded protein response (UPR) in eukaryotes involves splicing of an unconventional intron from the mRNA encoding the transcriptional activator of the pathway. In *Saccharomyces cerevisiae* a 252 nt unconventional intron is spliced out of the transcript of *HAC1,* changing the 3'end of the *HAC1* ORF and relieving the transcript from a translational block in a single step. The translational block is caused by base-pairing of part of the unconventional intron with the 5'UTR. In *Aspergillus niger* and other *Aspergilli* the unconventional intron in *hacA* mRNA is only 20 nt long. Since this intron is part of a stable stem-loop structure, base-pairing with the 5'UTR, in contrast to the yeast *HAC1* case is not possible. However, analysis of the *hacA* mRNA revealed a GC rich inverted repeat (18 base pairings). Upon activation of the UPR, the 5'UTR of the *hacA* mRNA is truncated by 230 nt, removing the left part of this inverted repeat. This implies a similar release of a translational block as in the case of the *S. cerevisiae HAC1,* but in two steps. The mechanism behind the 5'-truncation, which does not take place in either yeast *HAC1* or mammalian *XBP1* mRNA, has been unknown hitherto. Here we show that during secretion stress in *A. niger*, *hacA* transcription starts from a new start site closer to the ATG, relieving the transcript from translational attenuation. This transcriptional switch is mediated by HacA itself and the HacA binding site UPRE2 in the *hacA* promoter.

Introduction

In eukaryotic cells the majority of the secreted and membrane proteins are folded, assembled and modified in the endoplasmic reticulum (ER). Chaperones and foldases assist in folding the newly synthesised proteins [39], and correctly folded proteins are transported to the Golgi compartment where further modification takes place. The secretory proteins are transported in vesicles from the Golgi apparatus to the cell membrane where the content of the vesicle is released into the medium. When the amount of newly synthesised proteins exceeds the folding capacity of the ER, or if folding in the ER is impaired by other means, the eukaryotic cell triggers an intracellular signalling pathway in order to counteract the resulting secretion stress. This response is known as the unfolded protein response (UPR) [64,123,150,179], and leads to the increased synthesis of ER resident foldases and chaperones as well as many other components of the secretory pathway [160]. Genomic analysis of the UPR in *S. cerevisiae*, *Arabidopsis thaliana*, *A. nidulans*, *T. reesei*, and *A. niger* all linked large sets of genes (several hundreds) to the UPR, affecting all stages of the secretory pathway [4,48,89,154,160]. The central activation step of the UPR involves the splicing

of an unconventional intron from the mRNA encoding the transcriptional activator of the UPR. It is conserved from lower eukaryotes like yeasts, to higher eukaryotes like mammalian cells and mediated by Ire1p/IreA, a transmembrane kinase/RNase residing in the ER membrane. With its luminal domain Ire1p/IreA senses the 'folding state' of proteins in the ER, and when necessary a signal is transmitted over the ER membrane by activation of the RNase domain which in turn processes the messenger encoding the transcriptional activator of the UPR.

In *S. cerevisiae,* Ire1p-mediated splicing of a 252 nt unconventional intron from *HAC1* mRNA relieves the transcript from a translational block [66,137]. *HAC1* encodes a basic leucine-zipper type transcription factor that activates its target genes by a direct interaction with upstream activating sequences (UAS). In mammalian cells, the *HAC1* homologue *XBP1* is activated in a similar way by the splicing of a 26 nt unconventional intron [81], whereas the *Trichoderma reesei hac1*, the *Aspergillus nidulans* and *A. niger hacA* are activated by the splicing of a 20 nt unconventional intron [101,140]. To distinguish the unspliced *hacA* mRNA form the spliced form, they are referred to as $hacA^{u}$ (uninduced) and $hacA^{i}$ (induced) respectively.

Taking into account the magnitude of the UPR, and the risk that accumulation of unfolded proteins in the ER pose onto the cell, it seems plausible that natural selection has favoured strict controllability of the pathway. Although the UPR and its central activation step are conserved in all eukaryotic organisms, there are differences in the molecular mechanisms underlying the signalling pathway, that give the different organisms different ways of control over the response. In mammalian cells the UPR consists of three distinct pathways that operate in parallel, and include general attenuation of translation, apoptosis, and increased folding and degradation of unfolded proteins [50]. The mammalian UPR is triggered by at least three different ER stress sensors, all ER transmembrane proteins. This concerns: 1) IRE1, a nuclease which splices the unconventional intron from *XBP1* mRNA [186]. 2) ATF6, that migrates to the Golgi in response to ER stress. In the Golgi, its cytosolic domain, that encodes a transcription factor, is proteolytically released and relocated to the nucleus where it up-regulates its target genes [54]. 3) PERK, a protein kinase that in response to ER stress phosphorylates the translation initiation factor-2α, resulting in general inhibition of translation [52].

In contrast to mammalian cells, in *S. cerevisiae* the only known sensor of the UPR is Ire1p. In addition to the on/off switch of the yeast UPR, several other factors influencing the UPR circuit have been described. The transcription of *HAC1* was shown to be auto-regulated [112], and translation of the unspliced $HAC1^{u}$ mRNA blocked by long–range base pairing [137]. Furthermore, a second ER-to-nucleus signalling pathway was found, which leads to an increased production of Hac1p combined with production or activation of a putative UPR modulating factor [79]. Both signals are required for the super-UPR activation.

Finally, Gcn4p and its activator Gcn2p were found to be required for the induction of the majority of the UPR target genes [124]. Integrated with the Ire1p pathway, these elements turn the UPR from a simple switch into a more complex response, and give the yeast cell a more fine-tuned control over the UPR.

The UPR in *A. niger* shares some characteristics with the *S. cerevisiae* response, but there are some differences. As saprophytic organisms, filamentous fungi are capable of secreting large amounts of proteins to recruit nutrients. This characteristic might have shaped the secretory pathway and have set specific demands to the stress responses, which might explain the differences with the *S. cerevisiae* UPR. Firstly, in addition to the unconventional intron splicing, the 5'-untranslated region (5'UTR) of the *hacA* transcript of *A. niger*, *A. nidulans,* and *T. reesei* is truncated upon ER stress [101,140]. This has not been reported for the yeast *HAC1* or the mammalian *XBP1* homologues. Secondly, a novel type of transcriptional down regulation mechanism has been described in filamentous fungi, which was termed repression under secretion stress (RESS) [1,118]. Upon ER stress, it down regulates the transcription of some genes encoding secreted proteins, thereby lowering the burden on the ER. Thirdly, three UPREs are present in the *hacA* promoters of *A. niger, A. nidulans* and *A. fumigatus* [100]. These elements could play a role in the upregulation of *hacA* by the activator protein itself, which was experimentally shown in *A. niger* [101]. The position of the UPREs in the *hacA* promoter in relation to start sites of the full-length and truncated *hacA* mRNA could also implicate a role for them in the truncation of the *hacA* transcript.

Here we show the mechanism behind the truncation of the 5'UTR of the *A. niger hacA* mRNA, and the involvement of the UPREs in the regulation of the response.

Material and Methods

Strains and Transformations

A. niger AB4.1 (*cspA1, pyrG*), and transformants thereof were grown on minimal media plates containing per litre: 6 g $NaNO_3$, 1.5 g KH_2PO_4, 0.5 g $MgSO_4 \cdot 7\ H_2O$, 0.5 g KCl, trace elements [172], 1.5 % agar, and 20 g fructose as a carbon source (pH 6.0). 5 mM uridine was added in case of auxotrophy. Protoplasts were transformed according to standard procedures [76]. *Agrobacterium tumefaciens* mediated transformation of *A. niger* was performed according to previously described methods [25]. Transformation plates were incubated at 18^{o}C.

Construction of an *A. niger ΔhacA* Strain

A 4.5 kb fragment, containing *hacA* with 2 kb of its 5'flanking region and 1.1 kb 3'flanking region, was PCR amplified from *A. niger* genomic DNA using primers P1 and P2, and cloned into pCR-Blunt II-TOPO (Invitrogen). The resulting plasmid pHM58 was cut with *BglII*, filled-in with Klenow, and cut with *SalI.* This removed the major part of the *hacA* gene from the plasmid, which was replaced with the *hygB* cassette from pAN7-1 (accession number: Z32698). For this purpose, pAN7-1 was cut with *XbaI*, filled-in with Klenow and cut with *XhoI*. From the resulting plasmid, a 6,6 kb *KpnI-XhoI* fragment was isolated, containing the *hygB* cassette flanked by the upstream and downstream region of *hacA*. The fragment was cloned between the left- and right border of the *A. tumefaciens* plasmid pSDM14 [111]. This resulted in the plasmid pHM60, that was used for the replacement of *hacA* in *A. niger* by *A. tumefaciens* mediated transformation.

Construction Complementation Plasmid

A 3,2 kb *FspI-BamHI* fragment from plasmid pHM58 was ligated into the *HincII-BamHI* sites of pBluescriptSK$^+$, resulting in plasmid pHacA. For selective targeting of the plasmid to the *pyrG* locus of *A. niger*, the 3.9 kb *XbaI* fragment of pABpyrG*Not (kindly provided by Dr. Peter Punt, TNO Zeist, The Netherlands) was inserted into the *SpeI* site of pHacA, resulting in plasmid pHacA-pyrG*. The *pyrG** gene contains a filled-in *BamHI* site disrupting the ORF. Only homologous integration of the plasmid at the *pyrG* locus will restore the uridine prototrophy, enabling selective targeting of the construct to the *pyrG* locus.

Site-Directed Mutagenesis

For mutagenesis purposes of the elements present in the *hacA* promoter and 5'UTR of the *hacA* mRNA, a 623 bp fragment containing these elements was PCR amplified from pHacA with primer pair P3 and P4, and cloned into pCR-BluntII-TOPO (Invitrogen), resulting in pCR-Blunt-PhacA. Site-specific mutations were introduced into the various elements using the Quick Change site directed mutagenesis kit (Stratagene). The primers are listed in Table 1. The mutations were verified by sequencing performed with the Bigdye sequencing kit (Applied Biosystems) and analysed with the ABIPrism 3100 Genetic Analyser (Applied Biosystems). A 493 bp *BstEII-AspI* fragment of pHacA was replaced with the corresponding but mutated fragment from the pCR-Blunt-PhacA plasmid.

The unconventional intron was removed by digesting pHacA with *PstI* followed by re-ligation. The *hacA* variant with the unspliceable unconventional intron was constructed by ligation of the annealed oligonucleotide pair P10-P11 into the *PstI* digested pHacA plasmid. For mutation of the long range inverted repeats, a 1756 bp PCR fragment was

generated in several successive PCR and assembly steps, which involved primers P3, P4, P12, P13, P14, and P15 (Table 1). The resulting fragment was cloned into pCR-BluntII-TOPO, and after sequencing a 1351 bp *BstEII-BglII* fragment cloned into pHacA.
After introduction of the mutations in pHacA, the *XbaI pyrG** fragment was cloned into the *SpeI* site of pHacA. Constructs having the *pyrG** gene in the same orientation as the wild type pHacA-pyrG* plasmid were selected for transformation to *A. niger ΔhacA*.

EMSA

100 pmol oligonucleotide primers were annealed in 100 µl 10 mM TRIS-HCl pH 7.5 containing 100mM NaCl, by heating the mixtures to 95°C for 5 min and cooling them down to room temperature by placing the tubes on the work bench. 1 pmol of annealed oligo was end-labelled in a 20 µl reaction with 5 µl γ-^{32}ATP and 20 units T4 Polynucleotide Kinase (New England Biolabs). The reaction was incubated for 1h at 37°C, and the radio-labelled oligonucleotide purified on Biogel P30 spin columns (BioRad). Binding reactions (20 µl) were performed in the following buffer: 25 mM TRIS-HCl (pH 8.0), 5 mM dithiothreitol, 100 mM KCl, 4 mM spermidine, 0.1 µg/µl poly[d(I-C)], 0.25% bovine serum albumin, 5% glycerol, and included 1nM labelled probe, and 200 nM purified HacA protein [101]. Binding was allowed to proceed on ice for 15 min. The reactions were loaded on a 6% polyacrylamide gel and electrophoresed at 100 V in 0.25 x TBE (89 mM TRIS-HCl, 89 mM Boric acid, 20 mM EDTA, pH 8.3).

Southern Blot Analysis

Genomic DNA for Southern blot analysis was isolated by adding 800 µl buffer (100 mM Tris pH 8.0, 50 mM EDTA, 1 mM DTT, 35 mM SDS) to 0,1 g of ground mycelium. The DNA was purified by phenol-chloroform extraction according to standard procedures [141]. The DNA pellet was resuspended in H_2O and incubated with DNase free RNase (Roche Molecular Biochemicals). For each digest 10 µg of genomic DNA was incubated 6 h at 37°C with 20U of an appropriate restriction enzyme. The reaction products were separated on a 0.8% TBE agarose gel. Blotting and hybridisation was done according to standard procedures [141].

Table1. PCR oligonucleotides used in this study.

Oligo	Sequence	Remark
P1	5'-ggagaaccaaggtgctattg	-2071 bp *hacA* start codon
P2	5'-gtttaggagcctttacctcc	+1152 bp *hacA* stop codon
P3	5'-atttaccgtacggtcaattggggc	-503 bp *hacA* start codon
P4	5'-ccctgcctgtactgacgagtcatc	+120 bp *hacA* start codon
P5	5'-ggcctgatctga**a**cacg**gac**gccttttaaagagt	UPRE1
P6	5'-tcccccgttatgacacgg**ac**gcctgtgttcctgt	UPRE2
P7	5'-atggttcttaaggacacc**ac**tccttcttggccct	UPRE3
P8	5'-tcttttattgttct**c**tggttcttaaggacacc	ATG uORF
P9	5'-cgacctacatcaccgtcctccc**aacgtca**gcggttaagataagg ctcatagtaaatcgattg	5'splice-site
P10	5'-t**agt**gc**c**gt**cc**t**c**tgcga**t**ct**t**ca	unspliceable intron us
P11	5'-tg**a**aag**a**tcgca**g**a**gg**ac**g**gc**act**a	unspliceable intron ls
P12	5'-gcc**t**ggg**gt**t**a**gc**g**ccccc**t**gc**aag**ccccgttatgacacggtggcctgtgttc	Inverted repeat 5'UTR
P13	5'-gac**a**gg**t**aattcctg**c**ccccatgactttctcttcttcacagg	Inverted repeat *hacA* gene
P14	5'-tcccccgttatgacacggtggcctgtgttcctgt	-168 bp *hacA* start codon
P15	5'-tcaaaccgctcaagattcgttt	+1253 bp *hacA* start codon
P16	5'-cgtcgagaacgtcaaaggcgaacccgtc	For 5'RACE *hacA*
P17	5'-ttcgatgacaaggatgtccaga	For RT-PCR *bipA*
P18	5'-ggggatgagcttggtcatga	For RT-PCR *bipA*
P19	5'-atttaaataccctctcccatcgtcctc	For RT-PCR *pdiA*
P20	5'-cctcctcggcggtgcagtcaaccttcac	For RT-PCR *pdiA*

Mutations are represented in **boldface** type

Northern Blot Analysis

For northern blot analysis cultures were grown O/N at 30°C on minimal media supplemented with 0.2% w/v yeast extract. Aliquots of 0.5 g wet mycelium were transferred to 50 ml fresh medium and the cultures were incubated for 3 h before provoking ER-stress by the addition of either 20 µg/ml tunicamycin or 20 mM DTT. To the control cultures equal amounts of either DMSO of water was added. Samples from the stressed and unstressed cultures were taken 3 h after applying the stress agents. Since tunicamycin is a more specific agent to induce ER stress, northern blots of these treatments are shown where possible. For the analysis of the UPREs and *hacA* processing, the DTT blot is shown, as DTT resulted in much more pronounced *hacA* splicing.

For RNA isolation, mycelium was ground with 1mm glass beads using the Fast Prep FP120 (BIO 101 Savant), and total RNA was isolated using the RNeasy plant total RNA kit (Qiagen). In order to prevent degradation of the RNA during the purification procedure, 1.5 ml screw-cap tubes were filled with approximately 0.5 ml 1mm glass beads and 700 µl RLC buffer which was provided with the RNeasy plant total RNA kit. The filled tubes were frozen in liquid nitrogen prior to the harvesting of the mycelia. The mycelium was harvested rapidly and dried thoroughly between tissue paper, and a small piece of

mycelium (100-200 mg) was added to a pre-frozen tube and refrozen in liquid nitrogen. The frozen tubes were than placed in the Fast Prep FP120, and run on 4 cycles of 20 seconds on speed 6.5, which allowed thawing and simultaneous grinding of the mycelia. For Northern analysis 10 µg of total RNA was separated on a 1.5% (w/v) agarose gel containing 10 mM phosphate buffer (pH 7), and blotted onto a Hybond N membrane (Amersham Pharmacia Biotech). Blotting and hybridisation was done according to standard procedures [141]. Probes were generated by labelling DNA fragments with deoxycytidine 5'-[α-^{32}P]-triphosphate using the Prime-It II labelling kit (Stratagene), and purified on Biogel P30 spin columns (Biorad). Exposures were made on Curix HC-S Plus x-ray films (Agfa). Quantification of the northern blot signals was done by phosphor-imager analysis (Packard Instant Imager).

5'RACE and RT-PCR

For 5'RACE and reverse transcriptase PCR (RT-PCR) first strand cDNA was made from 1 µg of total RNA using the SMART RACE cDNA Amplification Kit (Clonetech). The 5´-end of the *hacA* transcript was amplified in a PCR with primer P16 using cycling conditions recommended by the kit's supplier. For the transcription analysis of *bipA* and *pdiA*, cDNA was amplified using primer pairs P17-P18 and P19-P20 respectively.

Microscopy

For microscopy, mycelia was grown between two glass slides in minimal media supplemented with 0.2% yeast extract, with and without the addition of 10 µg/ml tunicamycin. Images were taken using an Olympus BX51 microscope equipped with an Olympus DP11 digital camera.

Results

The *A. niger ΔhacA* Strain

To assess the role of the *hacA* gene and the control mechanisms in the UPR of *A. niger*, we aimed to obtain a corresponding knockout strain. Recently, *A. tumefaciens* mediated transformation has been shown to be an efficient tool for gene replacement in *Aspergillus awamori* [92]. Using this transformation procedure, the *hacA* gene of *A. niger* was successfully replaced with the hygromycin resistance cassette. The construct consisted of the

hygromycin cassette from pAN7-1 flanked by 2 kb of the *hacA* 5'-upstream and 1.5 kb of the 3'-downstream regions respectively.

Two types of colonies were present on the transformation plates. The majority of the colonies had a regular wild type phenotype, whereas approximately 30% consisted of small non-sporulating yellow colonies. Southern analyses performed on both type of colonies, showed the absence of the *hacA* gene in the aberrant colonies, whereas the construct was ectopically integrated in transformants having a wild type phenotype. Thus, deletion of *hacA* results in a distinct phenotype. Growth is severely impaired, and sporulation is almost absent compared to the wild type strain (Fig. 1A). On minimal media plates, as well as on plates with rich medium such as complete media (CM) and potato dextrose agar (PDA), the ***Δ****hacA* strain formed small compact colonies with only few conidia. The growth differences observed were analysed further microscopically. Under normal growth conditions, part of the hyphae of the ***Δ****hacA* strain are swollen compared to the wild type, and often these swollen hyphae are branched at the tips, forming finger-like extensions as shown in Figure 4B. In the presence of tunicamycin, growth of the ***Δ****hacA* strain was nearly abolished. The hyphae are shorter than in the wild type strain, and have swollen irregular shapes. Thus, application of ER-stress in the form of 10 μg/ml tunicamycin severely affected growth of the ***Δ****hacA* strain, indicating that a functional UPR is necessary for coping with ER-stress.

Northern analysis showed the inability of the ***Δ****hacA* strain to activate the UPR (Fig. 1B). In a wild type strain the up-regulation of the chaperone *bipA* and the foldases *cypB* and *pdiA,* due to ER-stress, coincides with the processing of *hacA* mRNA. In the ***Δ****hacA* strain, however, ER-stress does not trigger a comparable response. The basal transcript levels of *bipA*, *cypB* and *pdiA* are low and in response to tunicamycin treatment only a slight if any upregulation was observed. In contrast, retransformation of the ***Δ****hacA* strain with the wild type *hacA* gene fully restored the transcriptional pattern of regulation, typical for a wild type strain. The ***Δ****hacA* strain was used further as a background for studying the *hacA* gene expression mechanism.

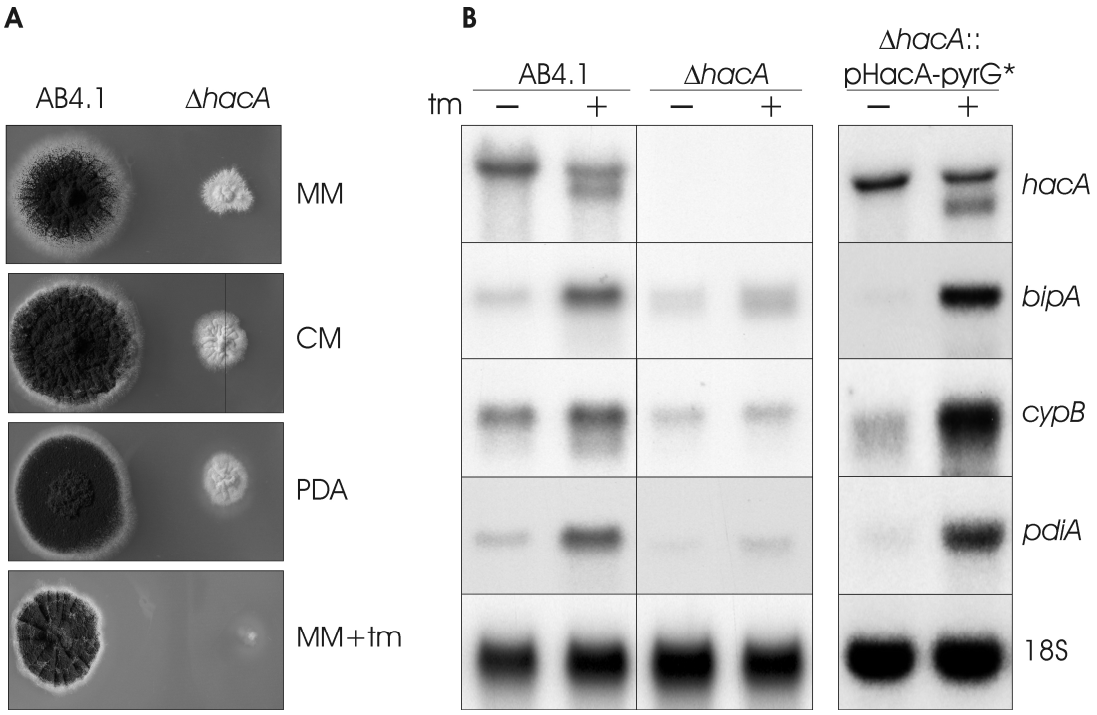


Figure 1. . Phenotypic and transcriptional analysis of the *A. niger ΔhacA* strain. **(A)** Growth of the *ΔhacA* strain compared to the wild type AB4.1 strain on different agar media. MM, minimal medium; CM, complete medium supplemented with 0.5% yeast extract; PDA, potato dextrose agar; tm, 10 µg/ml tunicamycin. **(B)** Northern blot analysis showing the lack of a functional UPR in the *ΔhacA* strain, and the restoration of the response by transforming the *ΔhacA* strain with plasmid pHacA-pyrG*. Probes that were used for hybridisations are shown to the right.

The Role of the Small ORF and the Putative 5'Splice-site

The 5'UTR of full-length transcript of *hacA* contains a small open reading frame (uORF) of 44 codons (Fig. 2A), which is not present in the truncated *hacA* mRNA. We previously hypothesised that this region could be involved in a translational control mechanism of *hacA*^{u} mRNA, since translational control is often associated with short ORFs, present in the 5'UTR of mRNAs [90]. The *S. cerevisiae* transcription factor Gcn4p and the mammalian transcription factor ATF4 are translationally controlled by small uORFs [51,59]. To test this hypothesis, the ATG start codon of the uORF was mutated to CTG, and the mutated *hacA* construct was introduced into the ***Δ****hacA* strain. However, northern analyses of the **Δ**uORF strain and subsequent signal quantification did not show clear effects of the mutation on the transcription levels of the foldases and/or the truncation of *hacA* mRNA upon application of ER-stress, when compared to wild type (Fig. 2B and C). It has to be noted that a second ATG is present in the reading frame of the 44-codon uORF of *hacA*. This ATG, however, is located 7 nt upstream of the central spacer nucleotide of UPRE2 (cf. Table I), and was therefore not mutated. As a result an eight-codon long uORF remains intact in the 5'UTR of the *hacA*^{u} mRNA of this strain.

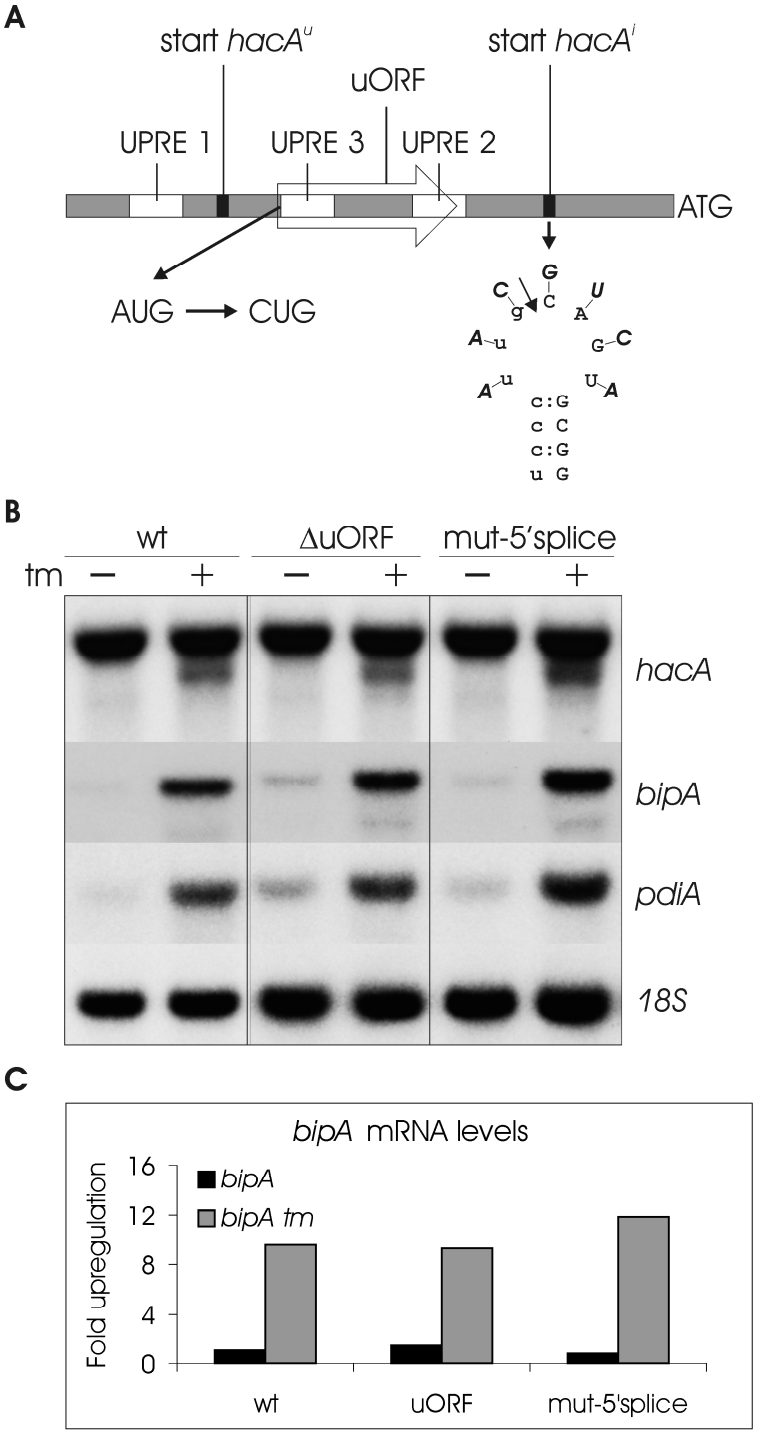


Figure 2. Mutational analysis of the uORF and the putative 5'splice-site, two elements present in the 5'UTR of the *hacA*^{u} mRNA. **(A)** Schematic overview of the 5'UTR of the *hacA* gene. ATG indicates the start codon of *hacA*, UPREs are indicated by open boxes, the start of *hacA*^{u} and *hacA*^{i} mRNA by black boxes, and the 44-codon uORF by an open arrow. The mutations introduced to disrupt the uORF and the putative 5'splice-site are indicated underneath the figure. The putative IreA cutting site is indicated by an arrow. **(B)** Northern blot analysis showing the effect of ER stress imposed by tunicamycin (tm) on the transcription of *hacA*, *bipA* and *pdiA*. 18S rRNA was as a loading control. **(C)** Relative *bipA* levels in the wt, the ΔuORF strain and the mut-5'splice-site strain in the absence (black bars) and presence (grey bars) of 20 µg/ml tunicamycin .

Interestingly, this uORF shows remarkable sequence identity with two small uORFs present in the *hacA* transcripts of *A. nidulans* and *A. fumigatus*. The translated sequences of these uORFs of *A. niger*, *A. nidulans* and *A. fumigatus* are MTRWPVFL, MTPWPLS, and MPTWLFL, respectively. Although this conservation could be solely the result of the fact that all three uORFs overlap with the UPRE2 sequence in each promoter, a function of the uORFs in translational attenuation cannot be excluded.

Previous inspection of the sequence around the start site of *hacA*^{i} mRNA revealed strong sequence similarity with the intron borders of the unconventional intron [101]. If this would lead to a similar secondary structure as is the case for the unconventional intron borders, this structure could be a potential target for IreA, and explain the mechanism behind the truncation of *hacA*. In contrast to the splicing of the unconventional intron which involves an IreA mediated cut in each of the two loops followed by re-ligation of the two exons by tRNA ligase [151], the hypothetical IreA-mediated 5'-truncation of *hacA* mRNA would involve only a cut at the recognition site. Although no RNA secondary structure incorporating the putative 5'splice-site in a stable stem-loop structure could be predicted, mutational analysis was performed to test whether or not this site would be a substrate for IreA. The seven non-paired nucleotides in the putative loop were mutated from purines to pyrimidines, or vice versa, in order to destroy a putative recognition site for IreA (Fig. 2A). If IreA indeed would be involved, this strategy should lead to the inability to form the truncated *hacA* messenger. However, northern analysis of a strain bearing such mutations, did not show any effect of the mutations in the putative 5'splice-site on the processing of the *hacA* mRNA (Fig. 2B). Three hours after tunicamycin-induced stress, the truncated *hacA* mRNA was present at levels comparable to the wild type strain. The up-regulation of *pdiA* and *bipA* mRNA levels upon stress was also similar to the wild type situation.

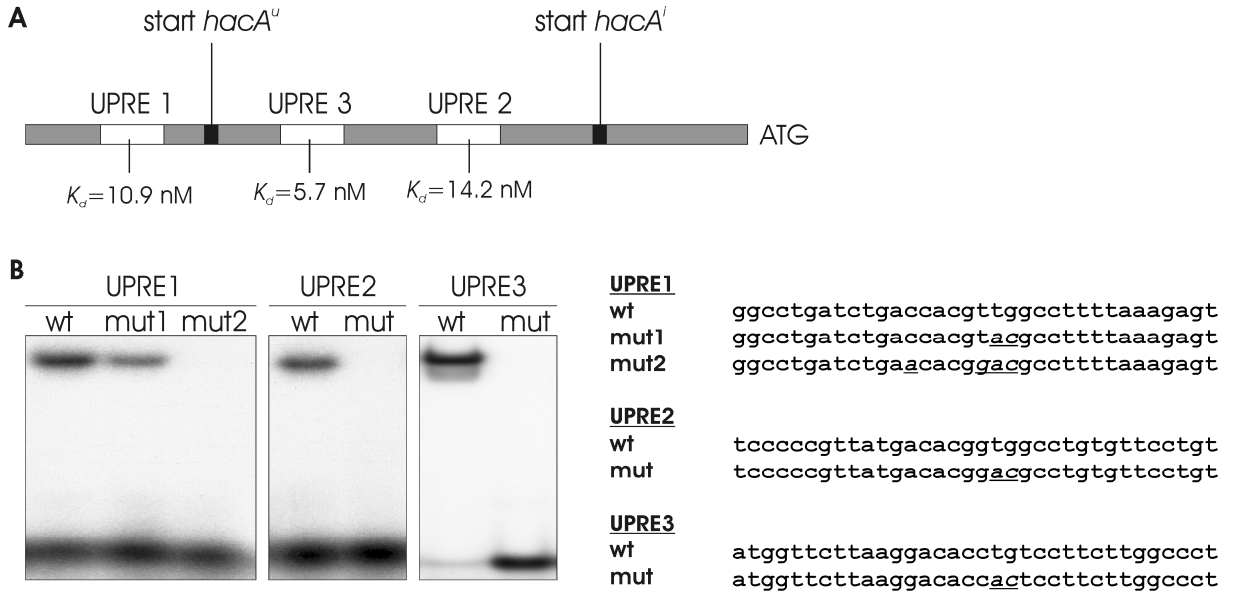
The Role of the UPREs in the *hacA* Promoter

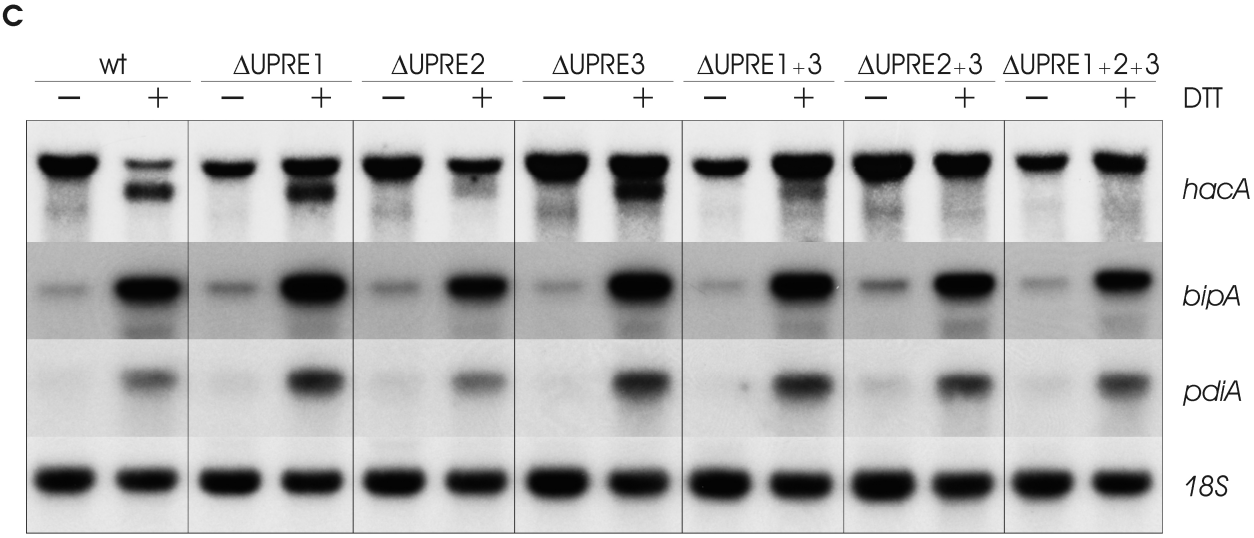
Previously, we identified by EMSA three UPREs in the *hacA* promoter. The presence and conservation of these three boxes in the *hacA* promoters of *A. niger, A. nidulans, and A. fumigatus*, suggests a function for all of them [100]. Mutational analysis was carried out to determine the contribution of each of these elements to the transcriptional regulation of *hacA*. Mutations were introduced in the core sequences of the *hacA* UPREs. The loss of binding due to the mutations was confirmed *in vitro* by EMSAs with HacA protein and 34 bp double stranded oligos comprising the different mutated UPREs. Whereas two mutations (Fig. 3A) were sufficient to completely abolish binding of HacA to UPRE2 and UPRE3, the corresponding 2 mutations introduced in UPRE1 did not lead to complete loss of binding. As a result, two additional mutations were introduced in the UPRE1. None of the strains carrying mutations in the *hacA* UPREs showed phenotypic growth changes on a

plate test (Fig. 4A), and these results were confirmed by light microscopy (data not shown). However, northern analyses performed on the strains revealed an altered transcriptional pattern for *hacA* and the foldases.

When UPRE2 was mutated separately, in combination with UPRE3, or in combination with UPRE1 and UPRE3, the truncation of *hacA* mRNA did not occur anymore in response to ER-stress (Fig. 3B). Although only the full-length transcript was formed in the UPRE2 mutants three hours after DTT treatment, *bipA* and *pdiA* transcripts were upregulated to the same extent as in the wild type strain, indicating that truncation of *hacA* is not a necessity for induction of the pathway. In contrast, mutation of UPRE1 or UPRE3 separately, did not lead to apparent changes in the splicing of *hacA* mRNA under ER-stress conditions. However, mutation of UPRE1 did result in lower basal levels of *hacA* mRNA (Fig. 3C). All three strains carrying mutations in UPRE1 had approximately half of the *hacA* mRNA levels compared to the wild type and the other mutant strains. This is in line with the finding that *hacA* is subject to auto-regulation [101]. Another observation is however more puzzling. In the wild type strain, the amount of the full-length *hacA* messenger had dropped significantly three hours after addition of DTT to the culture. This effect was also observed for the mutants **Δ**uORF, mut-5'splice and **Δ**IR (mutated Inverted Repeat, see below) upon DTT stress (data not shown). However, except for **Δ**UPRE2, this effect was not observed for the other UPRE mutants. On the contrary, full-length *hacA* levels were slightly induced by ER-stress in strains with either a mutation in UPRE1, UPRE3 or in both elements. This observation that UPRE1 seems to repress transcription of the full length *hacA* mRNA when HacA^{i} is abundantly present, would imply a dual function for UPRE1. Under normal conditions when the concentration of HacA^{i} is low, the UPRE1 functions as an activator site, whereas it functions as a repressor site during DTT stress. The underlying mechanism of this possible dual function of UPRE1 is unclear and will be subject to further investigation.

The transcription start sites of the three single UPRE mutants were determined by 5'RACE on RNA samples of non-induced and DTT induced cultures, to check whether mutations in the UPREs had an effect on the exact start site of *hacA^{u}* and *hacA^{i}* mRNA. Like in wild type, in all three mutants (**Δ**UPRE1, **Δ**UPRE2 and **Δ**UPRE3) the full length *hacA* started at 303 nt upstream of the ATG. In DTT stressed cultures of the **Δ**UPRE1 and **Δ**UPRE3 mutant, the transcription start site of the truncated *hacA* mRNA was found 73 nt upstream of the ATG, as in wild type.


A
start hacA^u
start hacA^i
UPRE 1
UPRE 3
UPRE 2
ATG
K_d=10.9 nM
K_d=5.7 nM
K_d=14.2 nM
B
UPRE1
UPRE2
UPRE3
wt mut1 mut2
wt mut
wt mut
UPRE1
wt ggcctgatctgaccacgttggccttttaaagagt
mut1 ggcctgatctgaccacgtacgccttttaaagagt
mut2 ggcctgatctgaacacggacgccttttaaagagt
UPRE2
wt tcccccgttatgacacggtggcctgtgttcctgt
mut tcccccgttatgacacggacgcctgtgttcctgt
UPRE3
wt atggttcttaaggacacctgtccttcttggccct
mut atggttcttaaggacaccactccttcttggccct


C
wt
ΔUPRE1
ΔUPRE2
ΔUPRE3
ΔUPRE1+3
ΔUPRE2+3
ΔUPRE1+2+3
− + − + − + − + − + − + − +
DTT
hacA
bipA
pdiA
18S

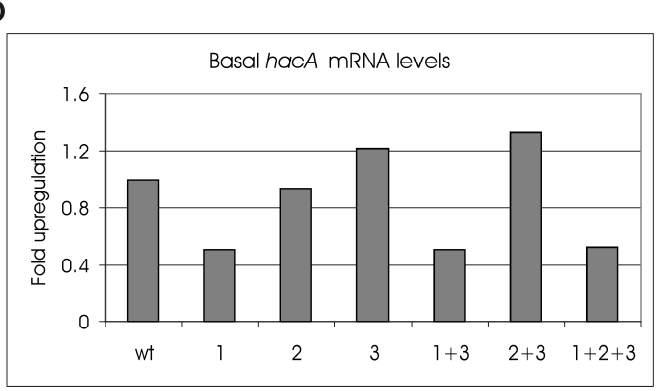

D
Basal hacA mRNA levels
Fold upregulation
1.6
1.2
0.8
0.4
0
wt
1
2
3
1+3
2+3
1+2+3

Figure 3. The role of the UPREs on the regulation of *hacA*. **(A)** Schematic overview of the promoter region of *hacA*. Indicated by black boxes are the transcription start points of *hacA*^{u} and *hacA*^{i}. Represented by open boxes are the UPREs. The binding constant of each UPRE is shown underneath the figure. **(B)** EMSAs showing the inability of HacA to bind to the mutated UPREs. The sequences of the 34 bp oligos representing the UPREs are shown, and mutations introduced in each UPRE to abolish binding by HacA are indicated in *italics* and are *underlined*. **(C)** Northern analysis showing the effect of mutated UPREs on transcription levels of *hacA*, *bipA* and *pdiA*. ER stress was imposed by the addition of 20 mM DTT to the media. **(D)** Relative *hacA* mRNA levels in the different UPRE mutants during non-stress conditions.

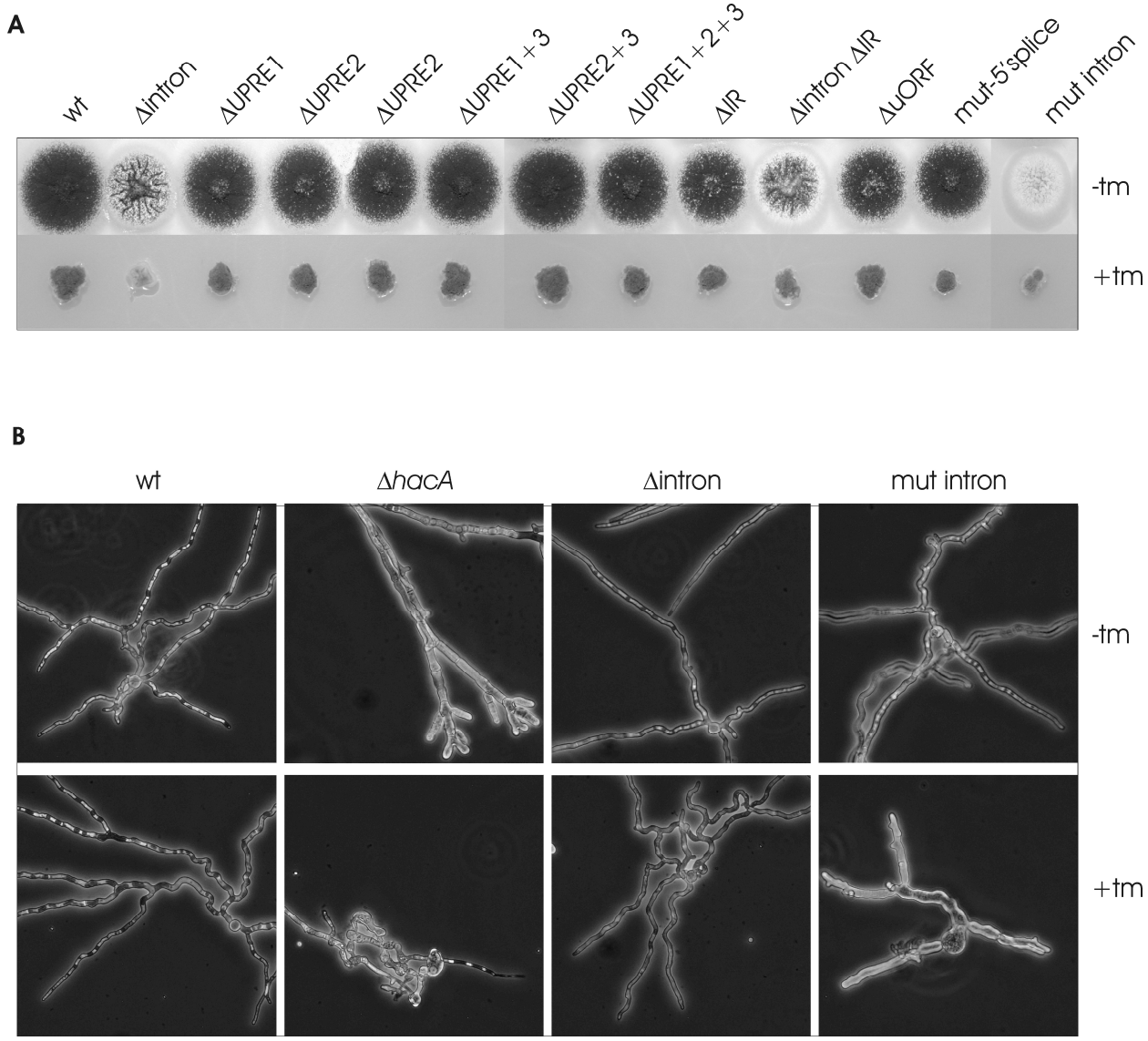


Figure 4. Phenotypes of the strains used in this study. **(A)** Growth of the wild type and mutant strains on complete (CM) media plates in the absence (upper row) or in the presence (lower row) of 10 µg/ml tunicamycin. **(B)** Microscope images of the wild type strain (wt), the *hacA* deletion strain (*ΔhacA*), the strain bearing the intron-less *hacA* gene (Δintron), and the strain bearing the unsplicable *hacA* variant (mut-intron). The bar indicates 1µm.

Long-range Base Pairing and Translational Block

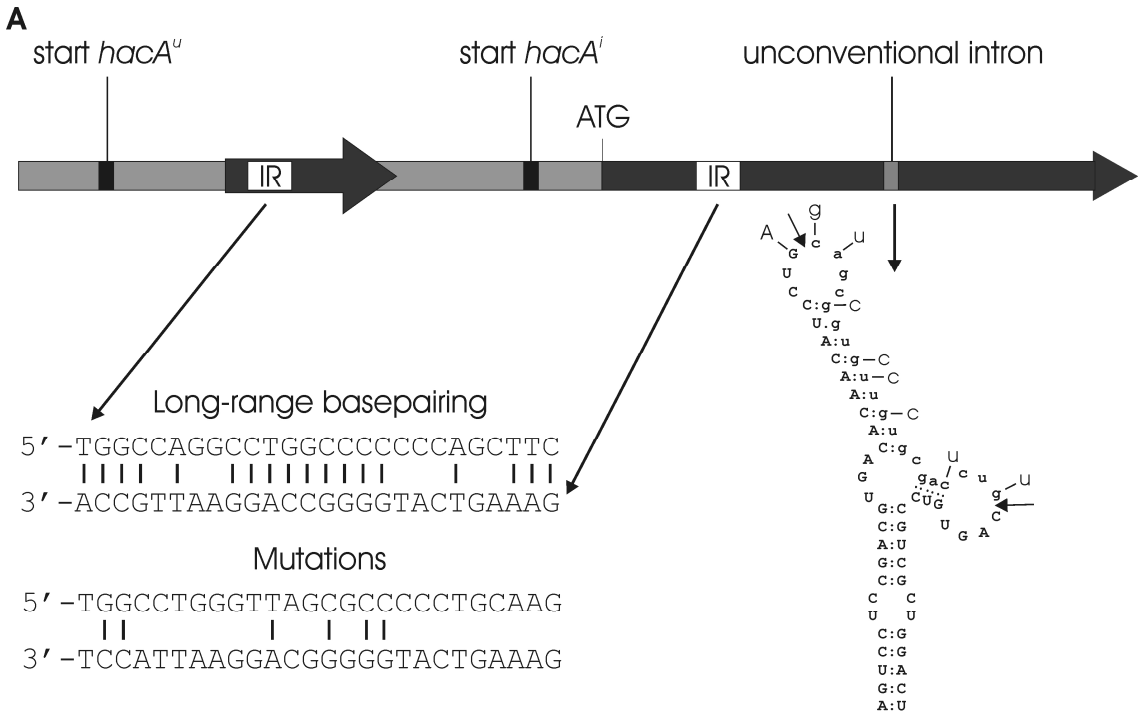
Long-range base pairing was shown to prevent the uninduced form of the yeast *HAC1* homologue from being efficiently translated [137]. Part of the 252 nt intron of the yeast *HAC1*^{u} mRNA forms a stable double-stranded structure by base pairing with the 5'UTR, thereby preventing the ribosomes from reading through. Unconventional intron splicing removes the 252 nt intron, and with it the right half of the inverted repeat, thereby releasing *HAC1*^{u} mRNA from its translation block.

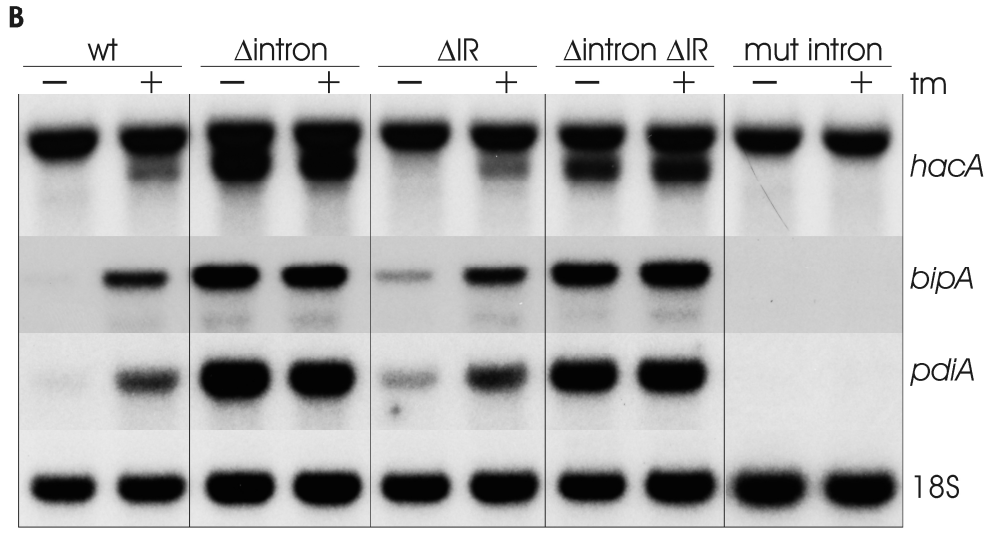
An identical mechanism is not possible for the *A. niger hacA*^{u} mRNA. The unconventional intron of *hacA* is only 20 nt long, and a stable stem-loop structure is predicted for this sequence, which leaves no nucleotides available for a putative long-range base pairing with the 5'UTR. However, sequence analyses revealed a 26 nt long inverted repeat in the *hacA*^{u} mRNA (Fig. 5A). This repeat contains 18 base pairs, of which 12 are GC pairs. In comparison, the yeast *HAC1* inverted repeat stretches out over 19 nt, and has 16 base pairs of which 11 are GC pairs. The left half of the inverted repeat in the *A. niger hacA* is located in the 44-codon uORF, and the right half in the *hacA* ORF close to the ATG. Truncation upon ER-stress removes approximately 230 nt of the 5'UTR of *hacA* mRNA including the uORF and thereby the left half of the inverted repeat, possibly releasing *hacA* from a translational block analogous to the yeast *HAC1* mechanism.

To determine the effect of the putative base pairing to the regulation of *hacA*, both half-sites of the inverted repeat were mutated. Since both half-sites are located in coding regions, only silent mutations were introduced to disrupt the putative base pairing (Fig. 5A). Although this strategy left 6 base pairs intact, it is likely that a stable base pairing of both half-sites is no longer possible. Analysis of the **Δ**IR strain by northern blots showed that under normal growth conditions approximately 3 times more *bipA* and *pdiA* transcript were present compared to wild-type (Fig. 5B), indicating higher basal levels of HacA in this strain. Under ER stress the transcriptional pattern of *hacA*, *bipA* and *pdiA* was similar as in wild type strain.

The Unconventional Intron

In order to study the unconventional intron splicing we aimed at constructing a ***Δire1A*** strain. However, multiple attempts and techniques to construct such a strain all failed presumably because of the lethality of the deletion. Therefore, the importance of the unconventional intron for the UPR was determined in the following two ways; by analysis of a strain carrying the *hacA* gene lacking the unconventional intron, and by a strain having a *hacA* variant with an unspliceable unconventional intron. Removal of the unconventional intron from the *hacA* gene resulted in a constitutive UPR (Fig. 5B).


A
start hacA^u
start hacA^i
unconventional intron
ATG
IR
IR
Long-range basepairing
5′-TGGCCAGGCCTGGCCCCCCCAGCTTC
3′-ACCGTTAAGGACCGGGGGTACTGAAAG
Mutations
5′-TGGCCTGGGTTAGCGCCCCCTGCAAG
3′-TCCATTAAGGACGGGGGTACTGAAAG


B
wt
Δintron
ΔIR
Δintron ΔIR
mut intron
tm
hacA
bipA
pdiA
18S

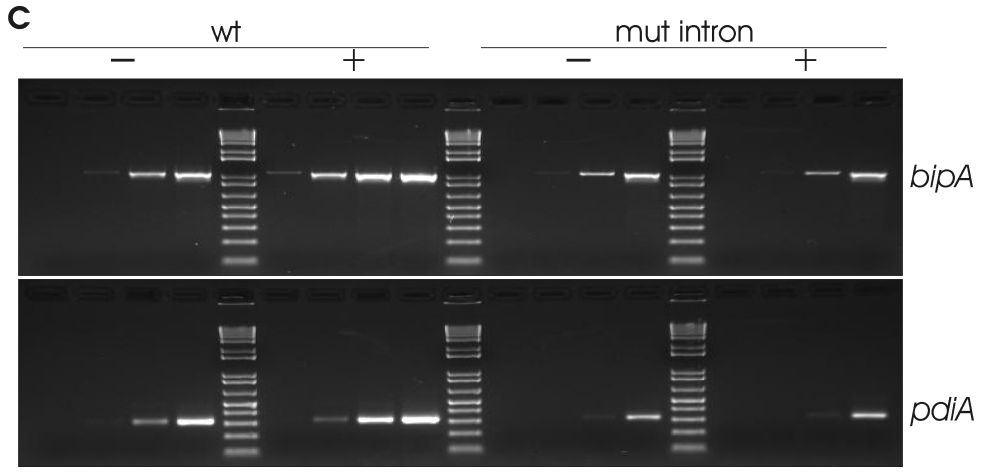

C
wt
mut intron
bipA
pdiA

Figure 5. The function of the unconventional intron, and long-range base pairing in the regulation of the UPR. **(A)** Schematic overview indicating the unconventional intron and the location of the inverted repeats in relation to other elements present in *hacA*. The *hacA* ORF is indicated by a solid black arrow, and the 5'UTR by a grey box. The inverted repeats (IR) are depicted by open boxes, and the base pairing of the inverted repeats underneath. The theoretically remaining base pairings after introduction of silent mutations is also indicated. The structure of the unconventional intron is depicted by the stem-loop structure, and mutations introduced to prevent splicing by disrupting the structure are indicated. The intron sequence is shown in lower case and cleavage sites indicated by arrows. **(B)** Northern blot analysis showing the effect of the different mutations or deletions on the expression and truncation of *hacA*, and on the transcript levels of *bipA* and *pdiA*. ER stress was imposed by treatment with tunicamycin, and the gene coding for 18S rRNA was used as control probe. **(C)** RT-PCR showing the transcript levels of *bipA* and *pdiA* in the wild type and the mut-intron strain. The 4 lanes in each condition correspond to 10, 15, 20, and 25 PCR cycles.

Both, the full-length and the truncated *hacA* transcripts were abundantly present under normal growth conditions, and the levels of *bipA* and *pdiA* were respectively 19-fold and 13-fold higher compared to the levels measured in the wild-type strain during non-stress conditions.

Induction of ER-stress by the addition of tunicamycin had no additional effect on the response in this strain, indicating that splicing of the unconventional intron is sufficient to fully induce the pathway. Northern blot analysis of a strain containing an unspliceable *hacA* variant (mut intron) revealed that splicing of the unconventional intron, and as a result HacA^{i}, is necessary to activate the downstream pathway (Fig. 5B). Truncation of *hacA* in response to tunicamycin induced stress and the accompanied upregulation of *bipA* and *pdiA* did not take place in this strain. *bipA* and *pdiA* transcripts could not even be detected on the northern blot in both conditions. It has been reported that deletion of *bipA* is lethal for *Aspergillus* [130]. Also, attempts to delete *pdiA* in *A. niger* were unsuccessful [105]. Therefore, the apparent absence of a *bipA* and *pdiA* transcript on the northern blot of the mut-intron strain is most likely a result of the limited exposure time of the X-ray film. To prove this, RT-PCR performed on RNA from the mut-intron strain which revealed that both *bipA* and *pdiA* transcripts were indeed present in this strain albeit at lower levels compared to the wild type strain (Fig. 5C).

Remarkably, both the **Δ**intron and mut-intron strain displayed clear phenotypes. When grown on agar plates the **Δ**intron strain produced fewer conidia than the wild type strain, whereas sporulation was almost absent in the mut-intron strain (Fig. 4A). Microscopic inspection of the mut-intron strain grown in the presence of tunicamycin

revealed a phenotype slightly comparable to the ***Δ****hacA* strain; poor growth and swollen hyphae (Fig. 4B).

Apparently, both HacA^{i} and HacA^{u} have physiological functions in *A. niger*, and the lack of either one results in the observed phenotype. The phenotype of the strain having the intron-less variant of the *hacA* gene, and therefore only capable of producing HacA^{i} protein, is not merely a result of permanent activation of the UPR. Over expression of *hacA^{i}* in a wild type background, leading to a strain which constitutively produces HacA^{i} protein but also capable of producing HacA^{u} protein from its endogenous *hacA* gene, did not lead to similar phenotypical changes (results not shown). This is in agreement with the previous observation that the amino acid sequence of the C-terminal part of the HacA^{u} from different *Aspergilli* is highly conserved [101]. The precise role of HacA^{u} in the fungus, however, is subject of further investigation.

Discussion

In eukaryotes, secretion stress leads to various responses which aim at homeostasis of ER functions. One of the secretion stress responses is the UPR, which is common to all eukaryotic cells. There are however differences in the molecular mechanisms underlying the UPR among different organisms. In this study, we describe some unique features of the *A. niger* UPR, not reported for the *S. cerevisiae* and mammalian UPR.

In both lower and higher eukaryotes, activation of the UPR starts with the splicing of an unconventional intron from the *hacA* transcript, mediated by IreA. Splicing of the 20 nt intron from the *A. niger hacA* mRNA is both necessary and sufficient to activate the UPR. A strain containing an unspliceable variant of *hacA* displayed poor growth and was unable to activate the pathway and up-regulate the foldases in response to stress, whereas a strain containing an intron-less variant of the *hacA* gene displayed a constitutive UPR (Fig. 5). In addition to the rather drastic on/off switch of unconventional intron splicing, other regulatory elements are present in the promoter of the *hacA* gene and in the *hacA^{u}* transcript, which provides the fungus additional regulatory control over the response.

Long-range base pairing in the *hacA* messenger prevents efficient translation of the full length transcript (Fig. 5). Opposed to the situation in *S. cerevisiae*, where pairing takes place between the 5'UTR and the unconventional intron, the 5'UTR of the *Aspergillus hacA* mRNA forms a double stranded structure with the coding region of the transcript, resulting in translational attenuation (Fig. 5). In the case of *A. niger*, the release of the translational block upon ER stress is exerted via a switch to an alternative transcription start located approximately 230 bp downstream of the usual transcription start site. This switch is

dependent on the UPRE2 and on HacA itself, thus confirming the auto-regulation of *hacA* transcription, demonstrated previously [101]. UPRE2 is the HacA target most proximal to the ATG and located 81 bp upstream of the start site of *hacA*^{i} mRNA. Mutation of UPRE2, either individually or in combination with UPRE1 and UPRE3, abolished the ability to synthesise the shorter transcript during ER stress (Fig. 3), indicating its importance for the transcriptional switch. The fact that the transcriptional switch depends on HacA itself was shown by two experiments. Firstly, a strain containing the *hacA* gene without the unconventional intron, and therefore only capable of producing the active form of the HacA protein, produced both the full length and the truncated *hacA* mRNA during normal growth conditions (Fig. 5). Secondly, a strain carrying an unspliceable version of the *hacA* gene (cf. mut-intron strain Fig. 5), and therefore unable to produce the active form of the HacA protein, was unable to synthesise the truncated messenger in response to stress.

Whereas a clear function can be assigned to the UPRE2, the roles of UPRE1 and UPRE3 in the regulation of *hacA* are somewhat more puzzling. On the whole, mutation of UPRE1 and UPRE3 individually or together left the response and the accompanying truncation of *hacA* intact. However, these sites could provide an additional fine tuning mechanism for the *hacA* transcription. When the UPRE1 was inactivated alone or in combination with other sites (**Δ**UPRE1, **Δ**UPRE1+3, and **Δ**UPRE1+2+3), the basal level of *hacA* transcription dropped to approximately 50% compared to wild type (Fig. 3C). In contrast, ER-stress led to a slight upregulation of the full-length *hacA* transcript in those same mutants, with the most pronounced effect in the double mutant **Δ**UPRE1+3 (Fig. 3B). Since the UPRE1 is most distal to the ATG of the *hacA* gene, and located 50 bp upstream of the transcription start point of *hacA*^{u} [100], it seems plausible that the UPRE1 affects the basal level of full-length *hacA* messenger. The basal level of *hacA* transcription is, however, not solely dependent on HacA. A strain unable to synthesise HacA^{i} (cf. mut intron in Fig. 5), displayed comparable *hacA* mRNA levels as the wild type strain under normal growth conditions.

Most likely, another, yet unknown factor plays an important role in the regulation of *hacA*. In mammalian cells, the *hacA* homologue *XBP1* is upregulated by ATF6 in response to ER stress [81], and in *S. cerevisiae,* certain conditions can boost the accumulation of *HAC1* mRNA via an Ire1p independent pathway [79]. A possible candidate involved in additional regulation of the *A. niger hacA* could be CpcA. CpcA is a homologue of the *S. cerevisiae* Gcn4p, which acts as the central transcription factor in response to amino acid starvation [177]. In addition, the *S. cerevisiae* Gcn4p plays an essential role in the UPR and is responsible for the induction of a major part of the UPR targets during ER stress [124]. Its mouse homologue ATF4 also has been shown to be involved in secretion stress [51]. The *T. reesei* and *A, nidulans* homologues *cpc1* and *cpcA* respectively, have also been linked to the UPR as they have been shown to be upregulated in response to ER stress [4]. Involvement of such a

factor could explain the slight upregulation of the full-length *hacA* transcript observed upon stress in the **Δ**UPRE1 mutant strains. The role of CpcA in the UPR and its effect, if any, on the *hacA* transcriptional regulation, perhaps via an interaction with the UPRE1 target, and the controversial role of the UPRE1 as an activator and a repressor site of *hacA* transcription remains to be established.

Whereas UPRE2 mediates transcriptional activation of *hacA* upon stress, an opposite function could be attributed to UPRE3. Mutation analysis of UPRE3 suggests that this site could act as a HacA scavenger in order to prevent the transcriptional switch from being activated at low stress levels, at which full activation of the pathway is not necessary. Disruption of UPRE3 resulted in higher levels of *hacA* both during normal and ER stress conditions. Of the three UPREs present in the *hacA* promoter, UPRE3 possesses the highest affinity for HacA, and stands apart from the other UPREs because of the missing spacer nucleotide between the two palindromic half-sites [100]. Using polysomal fractionation it has been shown that under non-stress conditions splicing of *hacA* still occurs, but at a low level [48]. The association of the spliced transcript with the polysomal fraction indicates that low levels of HacA^{i} protein will be synthesised during normal growth conditions, which is in agreement with our previous results from UPRE-promoter reporter constructs [100]. The UPRE3 could, therefore, have the function of a 'threshold' switch; only when sufficient HacA^{i} accumulates in the nucleus, the transcription of *hacA* will shift to the new transcription start site, possibly initiated via recognition of the UPRE2. Although further analysis is needed, we suggest a mechanism in which UPRE3 intercepts HacA^{i} when the protein comes into the vicinity of the promoter, thereby preventing activation of transcription via UPRE1 or UPRE2. Only when the HacA^{i} protein concentration reaches a certain threshold concentration, the pathway is further activated via the transcriptional switch to a new start of transcription followed by the consequent release from a translational attenuation check point.

We unravelled several mechanisms behind the *A. niger* UPR, that offers the fungus controllability over, and fine-tuning of the pathway. At the transcriptional level activation occurs in two steps: intron splicing followed by a transcriptional switch to a new start site. At the translational level, long-range base pairing inhibits efficient translation of the non-truncated *hacA* transcript. In addition to these mechanisms, our results indicate additional regulatory control of *hacA* transcription mediated via UPRE1 and HacA and most likely another transcription factor, possibly CpcA. UPRE3, finally, could function as a build-in safety mechanism to prevent unnecessary activation. The confirmation of these latter mechanisms and the integration into the regulation circuitry remains subject to further investigation.

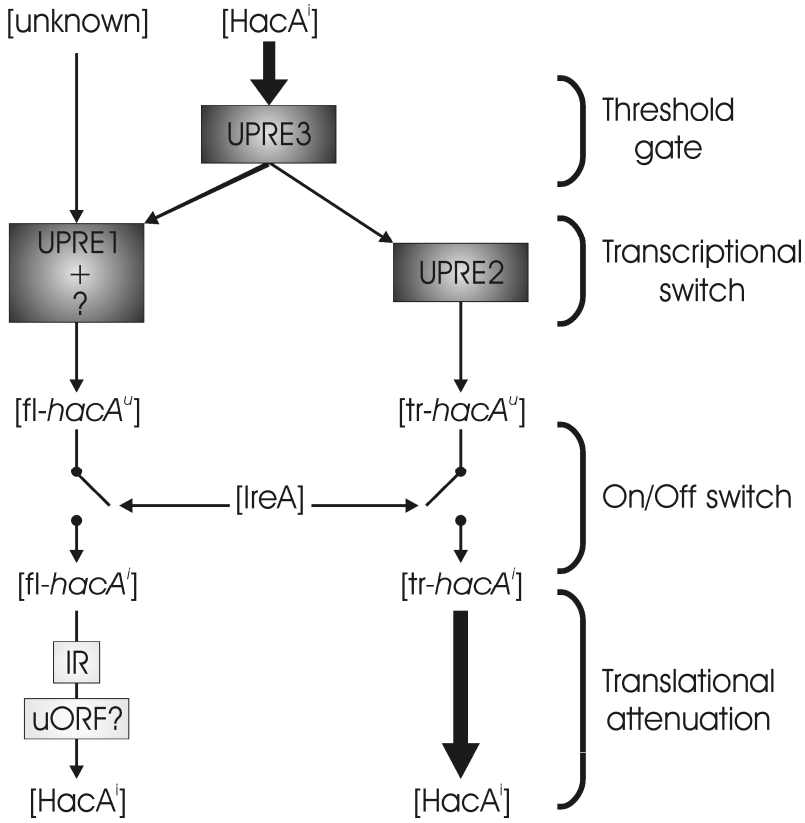


Figure 6. Schematic overview of the *hacA* regulation. The model depicts proteins and mRNAs between brackets indicating that their (active) concentration can vary. Elements present in the *hacA* promoter are in grey boxes, and elements present in the *hacA* mRNA in white boxes. The splicing of full length *hacA* (fl-*hacA*^{u}) and truncated *hacA* (tr-*hacA*^{u}) by IreA is depicted by the on/off switches. The translational attenuation is indicated by the thinner arrow compared to the non-attenuated translation.

Figure 6 shows a schematic overview of our current working model. It implies that under normal growth conditions low concentrations of active HacA^{i} protein will be present in the cell as a result of the limited intron splicing that constitutively occurs in the fungal cell. UPRE3 acts as a threshold gate, keeping HacA^{i} away from UPRE2 until the HacA^{i} concentration reaches a certain level. Transcription of *hacA* at this stage is mainly initiated via UPRE1 and/or an additional yet unknown promoter element, and is mediated by HacA^{i} and an unknown protein (possibly CpcA). The result is transcription of the full length *hacA*^{u} mRNA. IreA functions as the controller of an on/off switch to activate the UPR by splicing the *hacA*^{u} mRNA. The resulting full length *hacA*^{i} transcript is, however, subject to translational attenuation caused by long-range base pairing and possibly the uORF. As a result the production of 'active' HacA^{i} protein is slow at this stage. Only during prolonged stress will enough HacA^{i} protein accumulate to overcome the threshold value set by

UPRE3, and will the transcriptional switch be activated. The transcriptional switch is then operated by HacA^{i} that will bind to the UPRE2 and initiate its own transcription from a new start site. This will result in the truncated *hacA*^{u} transcript. This messenger lacks the translational block, and can be efficiently translated. However, it still contains the unconventional intron, and only after activation by IreA it will contribute to the active pool of HacA^{i} protein and only then the *hacA* dependent part of the UPR is at its full strength.

As a saprophytic organism, *A. niger* secretes large amounts of enzymes into its environment to degrade complex polymeric substrates used as natural nutrients for the fungus. This characteristic might also have set special demands on the secretion machinery of the organism and on the complexity of the regulation of the UPR in comparison with, for example, *S. cerevisiae*. Activation of the UPR affects a large set of genes [4,48,154], and unnecessary or unnecessarily high activation of the UPR will be costly in terms of energy. Therefore, the presence of several subsequent checkpoints allows the fungal cell to optimally tune its secretion machinery to the environment.

Chapter 5

General Discussion

General Discussion

Due to their natural capability to secrete large amounts of enzymes into their environment, filamentous fungi like *Aspergillus niger* and *Trichoderma reesei* are important protein production hosts for the industry. Production levels of natural enzymes from these organisms can reach up to 40 g/l and include amylases, cellulases, xylanases, lipases, proteases and oxidases. In contrast, production of heterologous proteins, especially those of non-fungal origin, is often hampered by low production titers, typically several orders of magnitude lower compared to homologous proteins. Although bottlenecks can exist at all stages of protein synthesis from transcription to secretion [47], the common assumption is that the main bottleneck is at the level of protein folding and maturation in the ER, since over expression of a heterologous protein is often accompanied by an upregulation of ER resident foldases and chaperones. This is a response of the cell to the presence of increased levels of unfolded proteins in the ER, and its function is to maintain homeostasis of the functions of the ER. The response involves the monitoring of the folding status of proteins in the ER and the transmission of the information to the nucleus in order to alter the transcription of relevant genes. In *A. niger,* IreA is the homologue of the yeast Ire1p and mammalian IRE1, which is the sensor of unfolded proteins in the ER. IreA transmits a signal across the ER membrane if the folding capacity of the ER is insufficient and unfolded proteins accumulate in the compartment. The signal is received by *hacA* which in turn transmits the signal to the nucleus. Genome wide transcription analyses have shown that the UPR affects a large set of genes [48,63,154]. Therefore, it can be expected that the activation of HacA, which is one of the key regulators of the UPR, is tightly regulated. This thesis describes the characterisation of the *A. niger* HacA and the regulation of its activation in response to ER stress.

Characterisation of HacA

The gene encoding HacA was cloned from *A. niger* var. *awamori*, and sequence analysis showed that it only shares limited homology with its yeast counterpart. Only the DNA binding domains of the proteins have been conserved. 30 of the 39 amino acids in the DNA binding domain are identical with the yeast Hac1p. Sequence analysis also revealed that the *hacA* gene possesses a rather small unconventional intron of 20 nt, in contrast to the 252 nt intron of the yeast *HAC1* gene. In spite of the low sequence identity at the amino acid level and the size difference of the unconventional intron, the activation mechanism by unconventional intron splicing of the messengers has been remarkably well conserved. As is the case for the yeast *HAC1*^{u} and the mammalian *XBP1*^{u} messenger, a stable stem-loop structure is predicted for the unconventional intron region of the *A. niger hacA*^{u} mRNA. The two intron-exon borders are present in loops of 7 nucleotides each, and the sequence

identity in these loops is very high among the three eukaryotes (see Fig. 3, Chapter 2), indicative of the high selection pressure on the conservation of the activation mechanism of the UPR during the course of evolution. Not unexpectedly therefore, a sequence alignment of the intron-exon borders of the *hacA/hac1* unconventional intron of several fungal species, shows that its sequence is almost identical in seven *Aspergilli*, *Fusarium graminearum*, *Magnaporthe grisea* and *T. reesei* (Fig. 1).

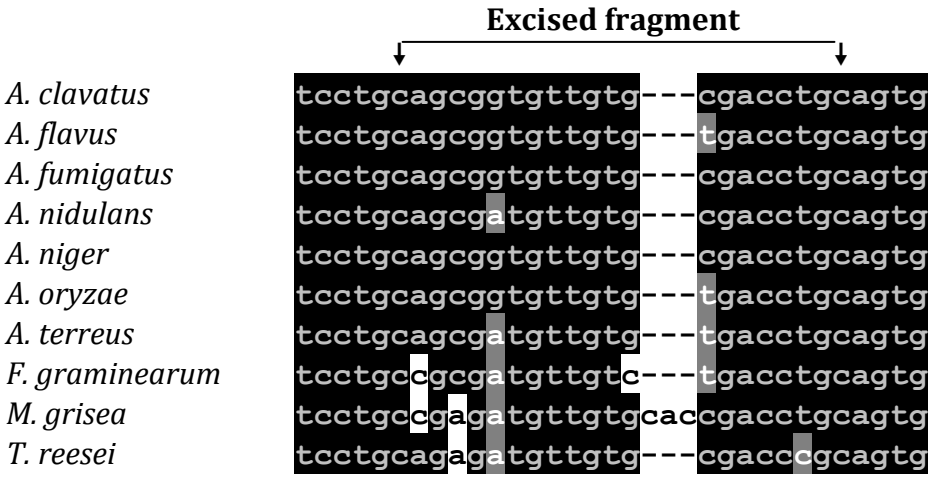


Figure 1. Alignment of the *hacA/hac1* unconventional introns from seven *Aspergilli*, plus *F. graminearum*, *M. grisea*, and *T. reesei*. The intron borders and the excised fragment are indicated by arrows.

Removal of the 20 nt unconventional intron from the *hacA* mRNA changes the reading frame of the transcript (Fig. 2). As a consequence a smaller protein is encoded in the spliced transcript. Whereas $hacA^{u}$ encodes a protein of 436 amino acids, $hacA^{i}$ encodes a protein of 342 amino acids. Both proteins share, however, the same N-terminal part containing the DNA binding domain and the leucine-zipper domain. This splicing-induced frameshift is similar to the activation mechanism of the mammalian *XBP1*. Splicing of the 26 nt intron from the mammalian *XBP1* also results in a frameshift. However, the translated protein from the spliced $XBP1^{i}$ transcript is significantly larger than from the unspliced $XBP1^{u}$ mRNA. In yeast, the unconventional intron splicing of *HAC1* does not result in a frameshift, but completely replaces the 3'-end of the transcript. Despite the differences, the unconventional intron splicing and, as a result, the encoded HacA protein shares some features with both the yeast *HAC1* and mammalian *XBP1* mechanism of activation.

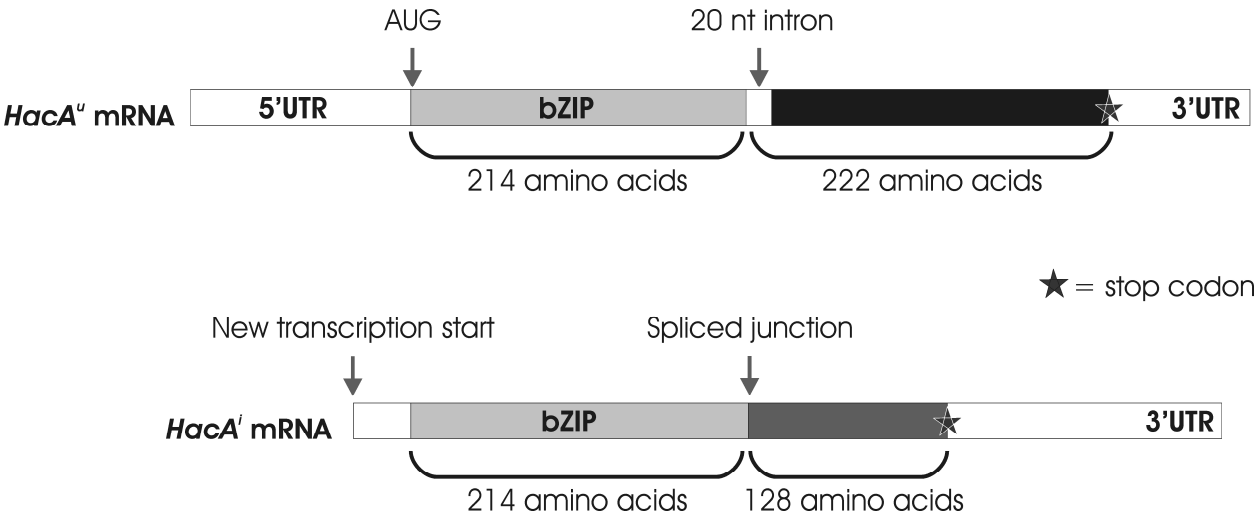


Figure 2. Unconventional intron splicing of the *A. niger hacA* transcript. The *hacA*^{u} mRNA contains an ORF encoding a 436 amino acid protein. Splicing of the 20 nt unconventional intron causes a frameshift, which results in a shorter ORF in the *hacA*^{i} transcript. The original 222 amino acids downstream of the bZIP domain in HacA^{u} are replaced by 128 amino acids in the HacA^{i} protein. In addition to the removal of the 20 nt intron, the *hacA*^{i} transcript in truncated at its 5'-end due to transcription from a new start site upon secretion stress.

Deletion of the *A. niger hacA*

Deletion of *hacA* in *A. niger* resulted in a severe phenotype, indicating the importance of *hacA* for the fungus (see Fig. 1, Chapter 4). The *ΔhacA* strain grows poorly on various media and hardly produces any conidia, making the handling of the strain difficult. In order to cope with secretion stress, *hacA* even appeared to be essential as growth was abolished upon tunicamycin treatment.

The severe phenotype of the *A. niger ΔhacA* strain is in contrast with the recently published *A. fumigatus ΔhacA* mutant. Although conidiation was decreased, the growth of the *A. fumigatus ΔhacA* strain was comparable to the wild type strain in the absence of ER stress [132]. The cause for this difference in phenotype between the *A. niger* and *A. fumigatus ΔhacA* mutants is not clear. It could for example indicate a distinct difference in the *hacA* dependent pathway between both *Aspergilli*, but might also be an artefact of the method of strain construction. Although different in the absence of secretion stress, both *ΔhacA* mutants share the sensitivity to ER stress. Growth was also found to be abolished in the *A. fumigatus ΔhacA* strain at concentrations of DTT and tunicamycin that could be tolerated by wild type.

The UPRE

HacA^{i} binds to UPREs present in the promoters of its target genes. The *Aspergillus* UPRE was identified as a partly palindromic sequence around a single spacer nucleotide, followed by a highly conserved CCT-stretch (Fig. 3).

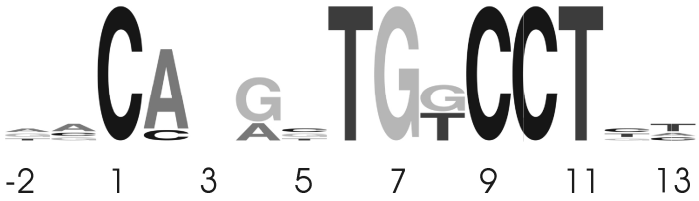


Figure 3. Sequence logo of the *A. niger* UPRE, based on an alignment of the UPREs identified in the promoters of *bipA, cypB, pdiA, prpA, tigA*, and *hacA*. The bases of the conserved core are numbered underneath the sequence, and the spacer nucleotide is located at position 4.

The UPREs are remarkably well conserved among homologous genes in various *Aspergilli*. For example, the core-sequence of the *prpA* UPREs of *A. niger*, *A. nidulans*, and *A. fumigatus* are identical (see Fig. 1, Chapter 3). Sequence alignment of larger promoter fragments shows that much more variation is present outside the UPREs (see Fig. 1, Chapter 3). So apparently there has been stringent selection on the preservation of the UPRE in homologous genes.

Since the various UPREs have different binding affinities for HacA^{i} (see Table 3, Chapter 3) the variation in the UPREs could be a mechanism to ensure differential upregulation of the various UPR target genes. The fact that a difference in binding affinity can result in differential expression was shown with a reporter gene expressed from the *cypB* promoter that had its original UPRE replaced by an artificial UPRE (see Fig. 5, Chapter 3). This element, obtained by binding site selection, possessed the highest binding affinity for HacA (see Table 3, Chapter 3), and when placed in the *cypB* promoter resulted in a higher expression of β-glucuronidase compared to the wild type UPRE.

The *hacA* promoter itself contains three UPREs, and the different affinities of these elements might have evolved to ensure the transcriptional switch mechanism as described in chapter 4. Sequence alignment of the *hacA* promoters of *A. nidulans* and *A. fumigatus* with the *A. niger hacA* promoter, showed that they also contain three UPREs, and that they are very similar (see Fig. 1, Chapter 3). Figure 4 shows an extended alignment of the UPREs identified in the *hacA/hac1* promoters of four additional *Aspergilli* species, *F. graminearum, M. grisea,* and *T. reesei*. In all promoters three putative UPREs were found with sequence identity to the *A. niger hacA* UPREs. As can be expected, the sequence identity among the *Aspergilli* UPREs is very high, whereas more variation is present when comparing the UPREs of evolutionary more distant fungal species to the *A. niger hacA*

UPREs. However, also in the *hacA/hac1* UPREs of *F. graminearum, M. grisea* and *T. reesei*, the majority of the residues of the *A. niger* consensus UPRE are present.

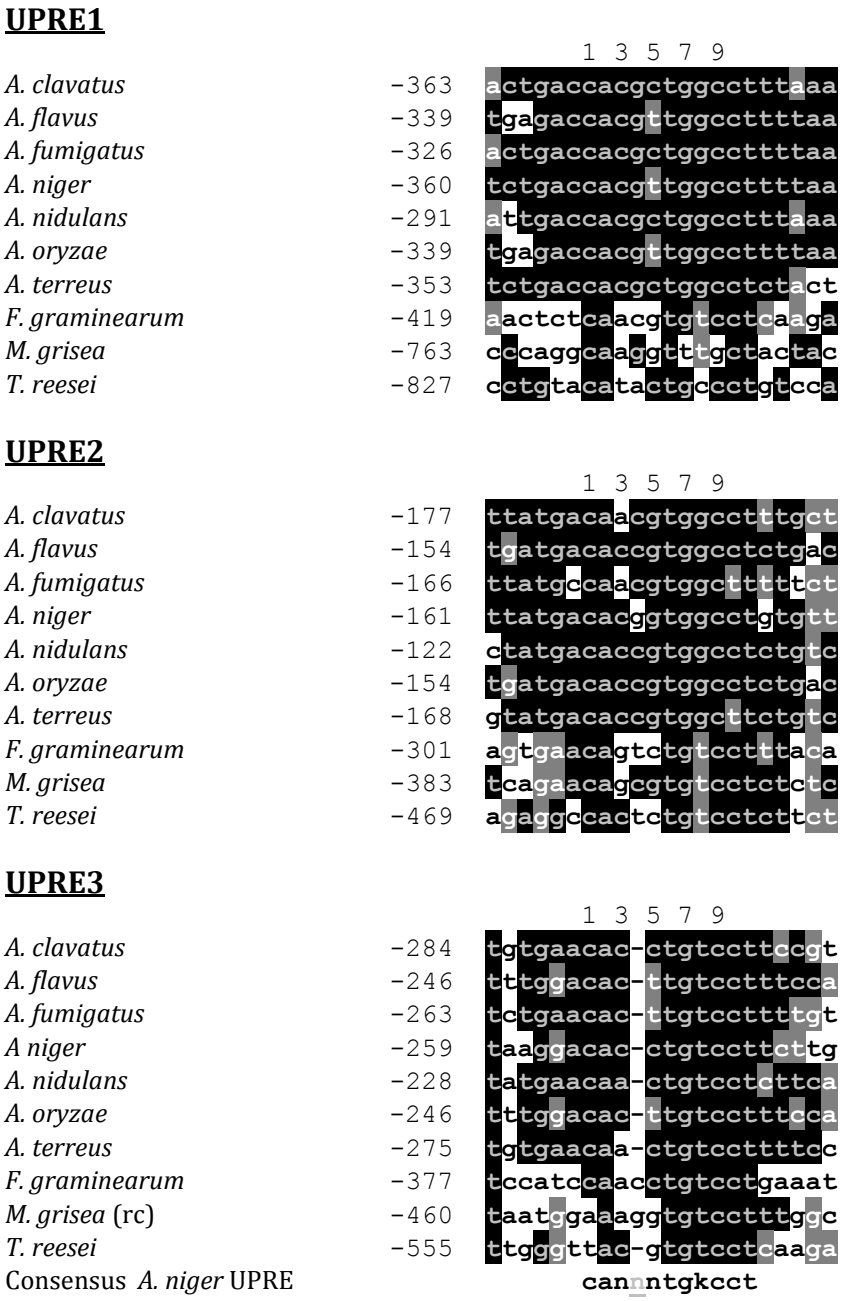


Figure 4. Sequence alignment of the UPREs identified in the *hacA/hac1* promoters of ten fungal species. The bases of the conserved core are numbered above the sequence. Underneath the sequence alignment the consensus sequence of the *A. niger* UPRE is given as a reference. The *M. grisea* UPRE3 is presented as reverse complement (rc). Numbers indicate the position of the UPRE relative to the ATG of the *hacA/hac1* ORF.

The Putative Activation Domain of HacA

In yeast, Ire1p mediated intron splicing removes a 252 nt intron that includes the coding part of the most C-terminal 10 amino acids of the Hac1p^{u} protein, and joins an 18 amino acid encoding part to the transcript (Fig. 7 intro). These 18 amino acids were found to function as an activation domain on Hac1p [98]. Sequence alignment of this activation domain from yeast with the *A. niger HacA^{i}* sequence revealed a region with some identity (Fig. 5), both being negatively charged.

Figure 5. Alignment of the 18 amino acid activation domain from the yeast Hac1p with amino acid 260-277 from the *A. niger* HacA.

To check whether unconventional intron splicing of *A. niger hacA* also joins an activation domain to the DNA binding domain, several truncated *hacA* proteins were expressed in the *ΔhacA* strain. All constructs were targeted to the pyrG locus using the plasmid pAB-pyrG* in the vector construction (Chapter 4). Since transcription of *hacA* is subject to auto-regulation, the truncated *hacA* variants were expressed from the *gpdA* promoter to allow for equal strong expression. Although not fully analysed, growth tests of the various transformants gave a hint that, at least, a part of the C-terminal domain of HacA^{i} is necessary for its activity (Fig. 6). A strain having the *hacA* gene truncated at the site of the unconventional intron and thus expressing HacA$_{1\text{-}214}$ was unable to grow in the presence of tunicamycin and displayed a similar phenotype as its parental *ΔhacA* strain, although it had slightly more conidia during normal growth. When truncated after amino acid 280, thus leaving the putative activation domain on the protein, still enabled a strain carrying such a construct to grow in the presence of tunicamycin. Under stress conditions, the phenotype of the strain expressing HacA$_{1\text{-}280}$ appears to be similar to wild type. Under normal growth conditions, however, this strain has a more 'woolly' appearance than wild type. It should be noted that the two truncated *hacA* variants lack the unconventional intron, and thus are unable to synthesise the HacA^{u} protein. The inability to produce HacA^{u} was earlier shown to result in a distinctive phenotype when grown on plates, by expressing an intron-less *hacA* gene from it endogenous promoter (see Fig. 4A, Chapter 4). The results presented here give an indication to the presence of an activation domain in the C-terminus of HacA^{i}, similar to the yeast *HAC1* activation domain, but more research is necessary to unveil the exact position, nature and function of the domain.

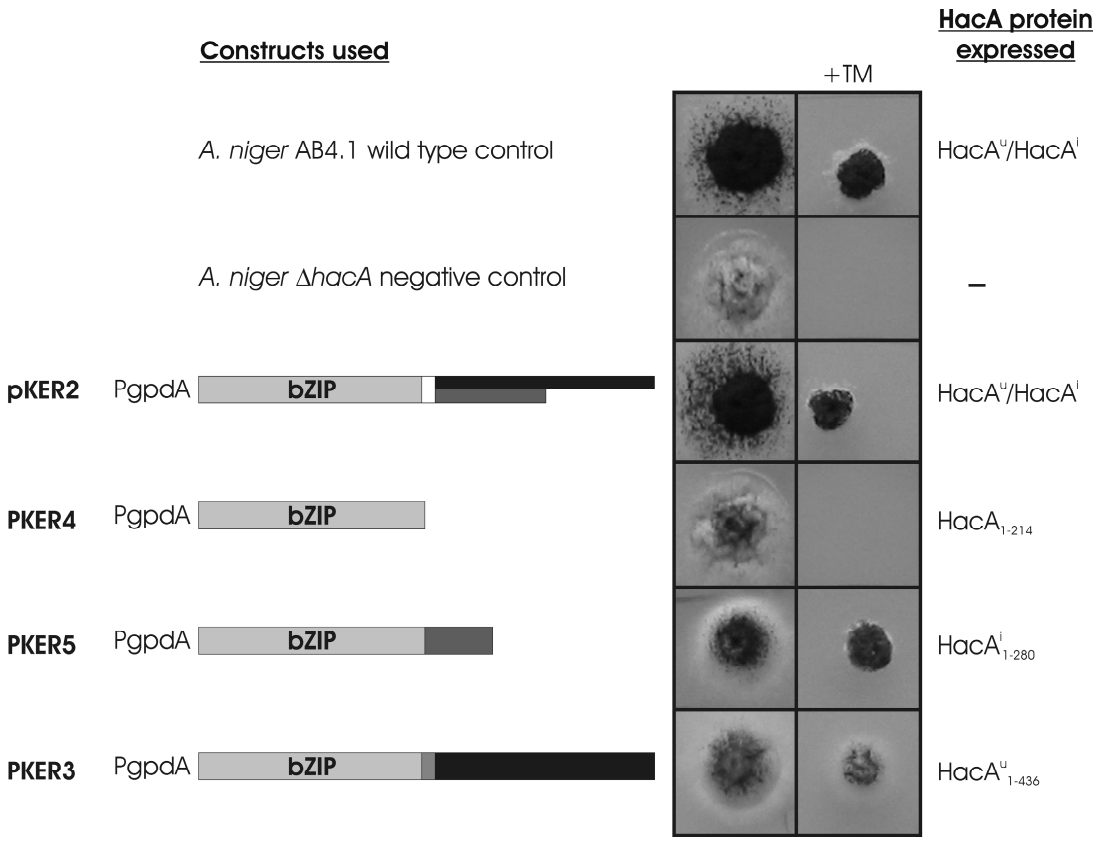


Figure 6. Comparison of the phenotypes during non-stress and secretion stress conditions of *A. niger* strains expressing various *hacA* constructs. A schematic overview of the genetic constructs used for strain construction is indicated on the left side of the figure. The light grey boxes indicate the Bzip domain that is identical in all constructs. The small white box indicates the 20-nt unconventional intron, the black boxes indicate the C-terminus of HacA^{u}. The dark grey boxes indicate the C-terminal domain of HacA^{i} .The small shaded box indicates the mutated intron due to which the transcript can not be spliced. TM indicates 10 µg/ml tunicamycin.

Expressing the *hacA* gene with the mutated and therefore unspliceable intron (see Fig. 5, Chapter 4 for mutations) from the *gpdA* promoter resulted in a similar phenotype on plates compared to a strain expressing the construct from its endogenous promoter (compare Fig. 6 and Fig. 4A, Chapter 4). Both strains show reduced sporulation on minimal media plates compared to the wild type strain, and are capable of growing in the presence of tunicamycin. The difference between the two strains is the lack of translational attenuation of *hacA^{u}* in the pKER3 strain. When expressed from the *gpdA* promoter, the *hacA* mRNA lacks the 3'-UTR and as a result can not form the translational block present in the native *hacA* messenger. As a consequence, more HacA^{u} protein might be produced in the pKER3 strain, but whether this has any measurable effects remains a subject to further investigation. However, the fact that both strains are capable of growth in the presence of tunicamycin, in contrast to the *ΔhacA* strain and the pKER4 strain, indicates that HacA^{u} indeed fulfils a function in the fungus and at least partly enables the fungus to overcome the lack of HacA^{i}.

Whereas these truncation experiments gave an indication that part of the C-terminal domain of HacA^{i} is dispensable for the fungus in coping with ER stress, sequence alignments revealed a conserved part at the end of the C-terminal domains of HacA^{i}/Hac1^{i} from ten fungal species (Fig. 7). This might indicate that also the most C-terminal part of HacA^{i} fulfils an important function (i.e. activation domain or binding site for other factors of the transcription initiation complex).

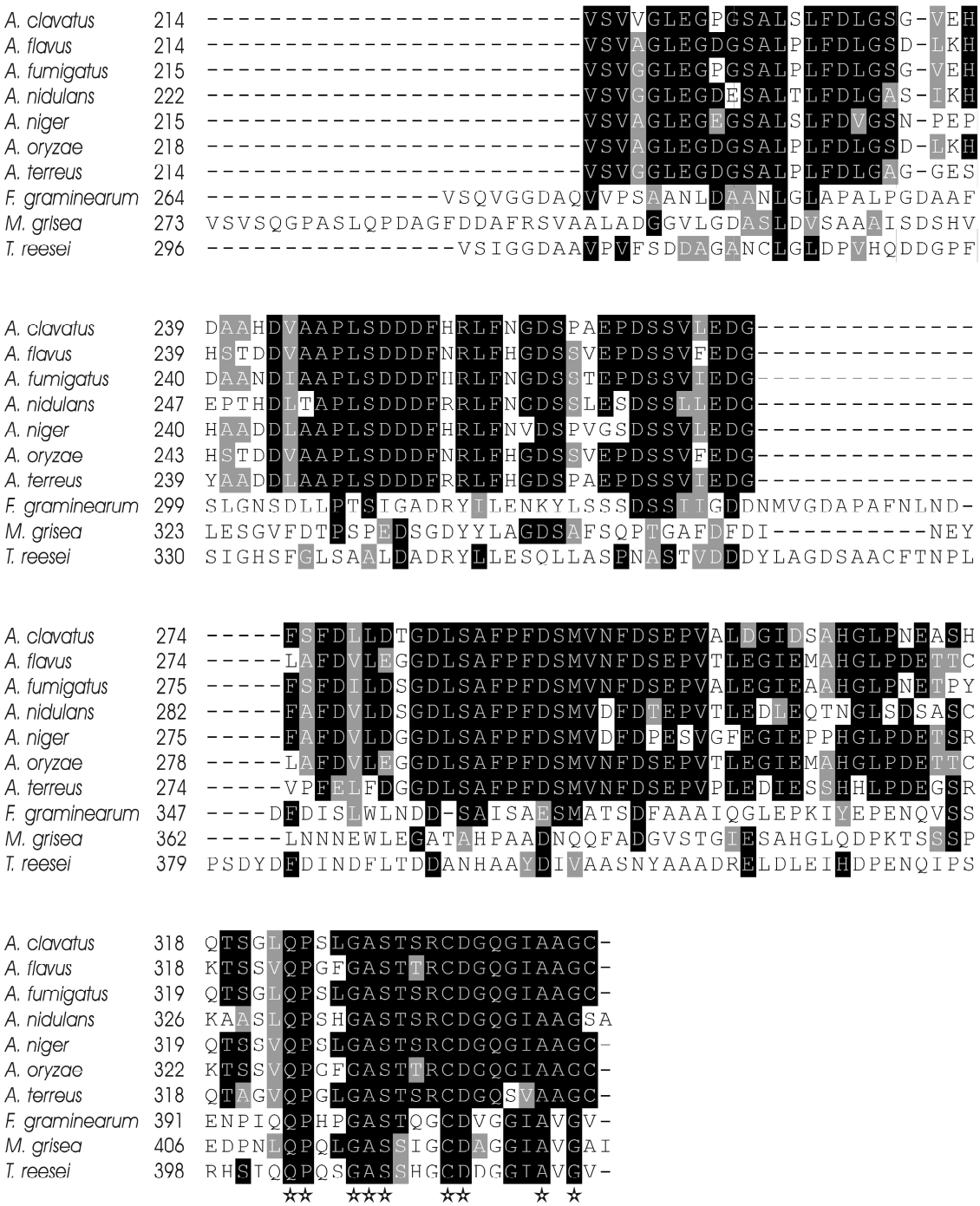


Figure 7. Alignment of the amino acid sequences of the C-terminal domains of HacA^{i}/Hac1^{i} from ten fungal species. The conserved residues at the C-terminus of the ten HacA/Hac1 proteins are indicated by asterixes.

The C-terminal Domain of HacA^u

The C-terminal domains of HacA^u from *A. niger*, *A. nidulans* and *A. fumigatus* are highly conserved, even more than the C-terminal domains of the induced HacA^i proteins (see Fig. 4, Chapter 2). This could be due to a specialised yet unknown function that HacA^u has in the fungus. An indication that HacA^u indeed fulfils a function in the cell comes from a strain carrying an intron-less version of the *hacA* gene (see Fig. 4A, Chapter 4). This strain cannot produce HacA^u, and although it can cope with tunicamycin stress, its growth during non-stress conditions is slightly impaired and it generates less conidia under both conditions compared to wild type. It is unknown whether significant levels of HacA^u protein will be produced in *A. niger*, but a possible function for HacA^u would be the role of a negative regulator of the UPR. In mammalian cells it was recently shown that XBP1^u acts as a negative regulator of XBP1^i and ATF6 [189].

Similar to the splicing of *hacA*, splicing of the 26 nt intron from *XBP1* changes the reading frame of the transcript downstream of the DNA binding domain and dimerisation domain, resulting in the switch of a large part of the C-terminal domain. It was found that XBP1^u shuttles between the nucleus and the cytoplasm due to a nuclear exclusion signal in its C-terminal domain. XBP1^u targets both XBP1^i and ATF6 in the nucleus by associating with them, and targets them subsequently to the proteasome for degradation. Whether such a function also resides in the C-terminus of HacA^u remains to be examined, but sequence analysis of the C-terminal domain of HacA^u showed that it is a leucine rich domain. Of the 222 amino acids of the C-terminal HacA^u domain 40 are leucine residues, accounting for 18% of the domain. In comparison, the C-terminal domain of HacA^i has only 11 leucine residues. This could be indicative of the presence of a nuclear exclusion signal (NES) in the C-terminus of HacA^u. NES's are leucine rich nuclear export signals that in yeast have been shown to interact with the Crm1p (Chromosome Region Maintenance protein 1), a nuclear exportin [33]. In addition, a sequence alignment of the C-terminal domains of ten fungal HacA^u/Hac1^u proteins, showed that they all are leucine rich (Fig. 8).

A NES is indeed predicted in the C-terminus of *A. niger* HacA^u (L245, L247, L253) by the NetNES 1.1 server (http://www.cbs.dtu.dk/services/NetNES/) [77], whereas none can be found in HacA^i. Thus HacA^u might well play a role similar to the mammalian XBP1^u, and shuttle between the nucleus and cytoplasm in order to transport HacA^i to the proteasome for degradation. However, this can not be the only function of HacA^u, as (over) expression of HacA^u in the absence of HacA^i partly rescues the phenotype of the *ΔhacA* strain (Fig. 1 and 4, Chapter 4 and Fig. 6). These strains express *hacA* with a mutated unconventional intron that can not be spliced, and thus can only produce HacA^u protein. Either HacA^u directly interacts with promoter elements and alters transcription of downstream genes, or it facilitates the nuclear export (besides HacA^i) of other bZIP transcription factors for degradation.

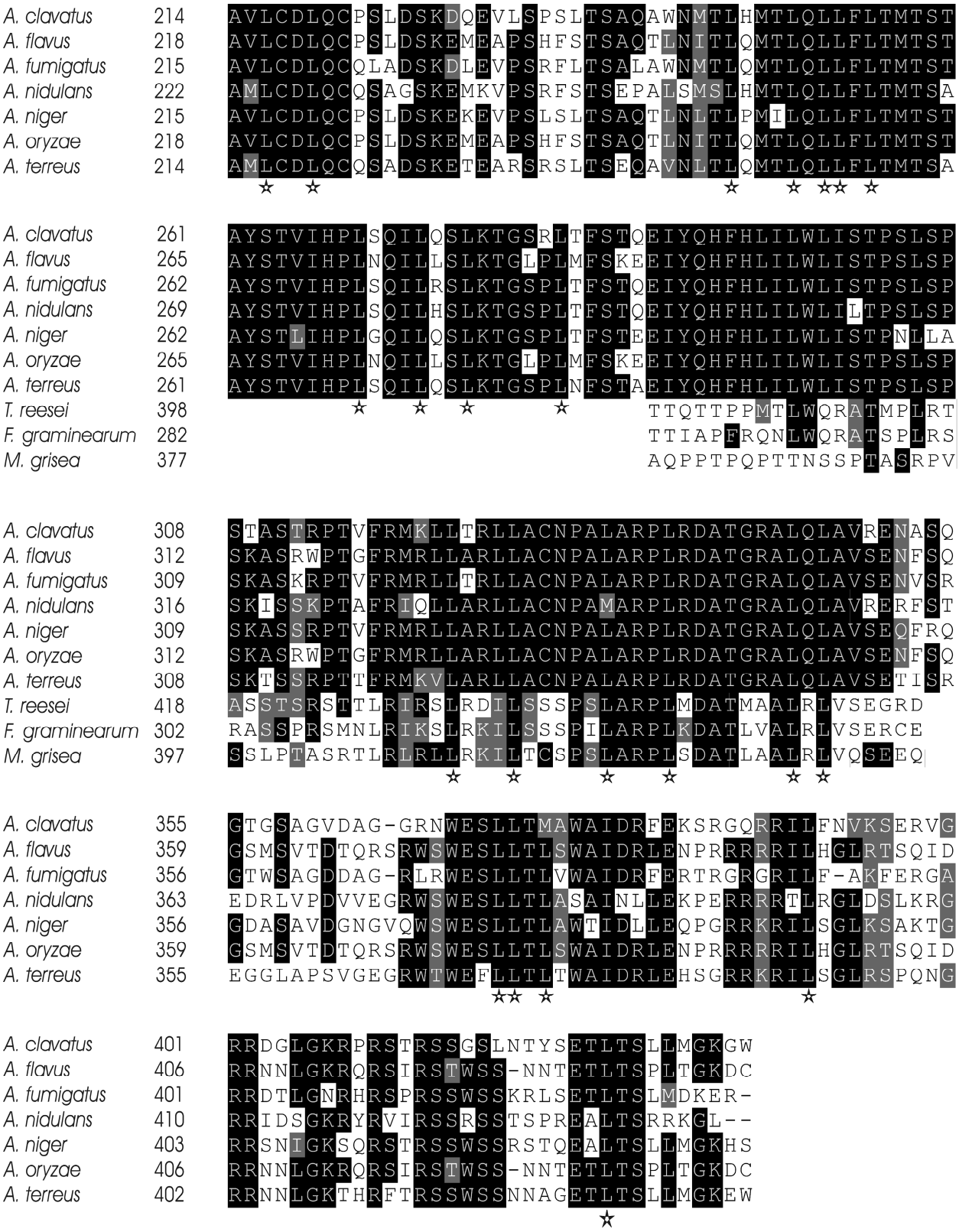


Figure 8. Alignment of the amino acid sequences of the (partial) C-terminal domains of HacA^{u}/Hac1^{u} from ten fungal species. The conserved leucine residues are indicated by asterixes, which might be indicative for the presence of a NES in these domains. From *T. reesei, F. graminearum*, and *M. grisea* only that part of their Hac1^{u}/HacA^{u} proteins is shown that exhibits significant homology with the *Aspergilli* HacA^{u} proteins.

Long-Range Base Pairing and the Transcriptional Switch

In addition to the unconventional intron splicing which is a common feature among eukaryotes, the *hacA*^{i} transcript is 230 nt shortened at its 5'UTR compared to the full length *hacA*^{u} transcript. This is the result of the synthesis of *hacA* from a new transcription start, closer to the start codon of *hacA*, and is mediated by HacA^{i} itself and via UPRE2. As a result of this new transcription, four different *hacA* transcripts can be present in the fungal cell: full length unspliced *hacA*, full length spliced *hacA*, truncated unspliced *hacA* and truncated spliced *hacA.* This has some implications.

Due to long range base-pairing, the full length *hacA* transcript is subject to translational attenuation, similar to yeast *HAC1*^{u} (see Fig. 8, Chapter 1) [137]. Contrary to yeast, splicing of the intron does not relieve the *hacA* transcript from the translational block (Fig. 9). As a consequence, the initial IreA mediated splicing of *hacA* mRNA does not immediately turn on the pathway at its full strength. The full length *hacA*^{i} transcript has first to escape the translational block and yield sufficient HacA^{i} protein that subsequently can operate the transcriptional switch. The newly synthesised truncated *hacA*^{u} mRNA then needs to be spliced by IreA before it can contribute to the HacA^{i} pool. Thus, the uncoupling of the translational block from the unconventional intron splicing, offers *A. niger* an additional level of control in activation of the pathway compared to yeast.

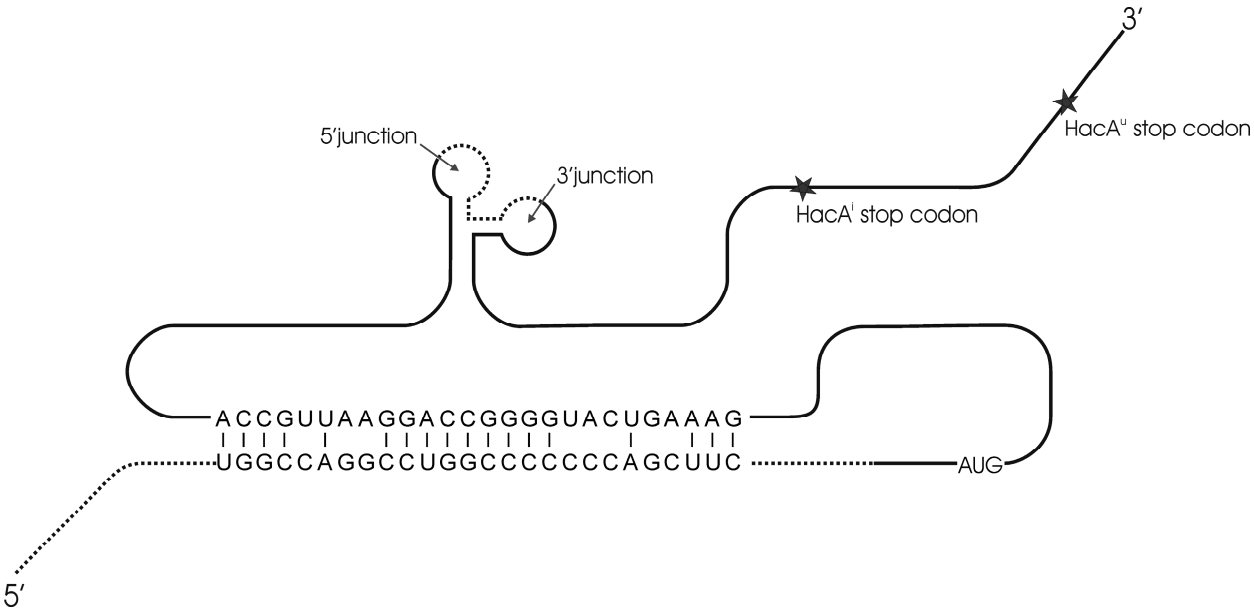


Figure 9. Schematic representation of the secondary structure and long range base pairing in the *A. niger hacA*^{u} mRNA, causing translational attenuation. The *hacA* 5'UTR and *hacA* coding sequence contain complementary sequences. Shown as dashed lines are the unconventional intron, and the part of the 5'UTR that is lost from the *hacA* transcript upon secretion stress. The intron borders are indicated by the small loops, en the splice sites are indicated by the arrows. The stop codons present in the reading frame of *hacA*^{u} and *hacA*^{i} are indicated by asterixes.

If HacA^{u} indeed would function as a negative regulator of the UPR as speculated earlier, this two step activation of the UPR would also offer the fungus an ingenious way of quickly shutting the pathway off. Ceasing the IreA dependent splicing would leave a pool of truncated but unspliced *hacA* mRNA in the cell. Not subjected to translational attenuation, this could then readily be translated into HacA^{u} protein and target the HacA^{i} pool for degradation, thereby quickly stopping unnecessary activation of target genes.

CpcA

In *S. cerevisiae* Gcn4p plays an important role in the regulation of the UPR (see Fig. 9, Chapter 1). Previously we speculated on a role for CpcA in the transcriptional regulation of *hacA* (see Fig. 6, Chapter 4). CpcA is the *A. niger* homologue of the yeast Gcn4p, and as a transcriptional activator might be responsible for the transcription of the full length *hacA*, even when UPRE1 is mutated (see Fig. 3C, Chapter 4). A sequence alignment of *hacA* promoter fragments from seven *Aspergilli* revealed a highly conserved sequence stretch just upstream of the UPRE1, which may indicate a binding site for a yet unknown regulator (Fig. 10). Interestingly, the border of this conserved box shows sequence similarity with the CpcA/Gcn4p binding site, TGACTC, which could be clue to the involvement of CpcA in the *Aspergillus* UPR.

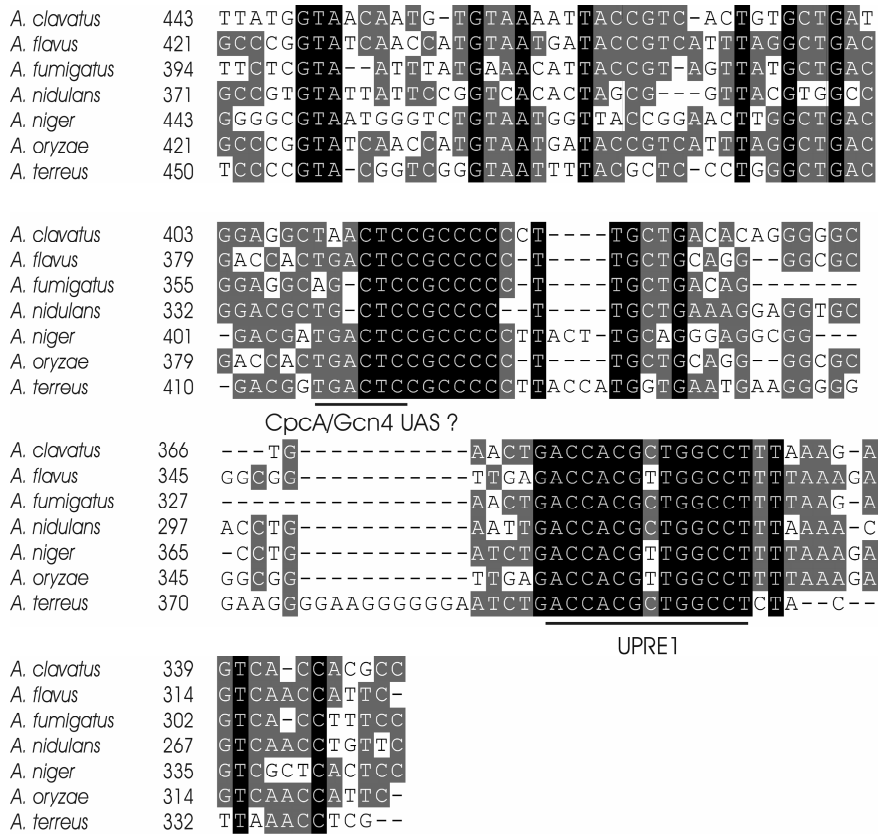


Figure 10. Alignment of *hacA* promoter fragments from seven *Aspergilli*. The conserved *hacA* UPRE1 box is underlined. Upstream of the UPRE1 is another conserved sequence stretch. The putative CpcA binding site is underlined.

Summary

Aspergillus niger is a filamentous fungus that is ubiquitous and commonly found on decaying plant material. *A. niger* has a saprophytic lifestyle and plays an important role in the degradation and recycling of dead plant material. In order to break down the complex plant polymers into smaller molecules that can be taken up and serve as energy and nutrient sources, the fungus produces a variety of hydrolytic enzymes. Many of the secreted enzymes of *A. niger*, like amylases, proteases, pectinases, xylanases andlipases find applications in the baking, starch, textile, food and feed industry. In industrial fermentations these extracellular enzymes can be produced in large quantities, some up till 40g/l. Although the high capacity secretory pathway of *A. niger* is often successfully exploited for the production of homologous and other fungal enzymes, the expression of heterologous enzymes of non-fungal origin is often several orders of magnitude lower. It is generally accepted that protein folding in the endoplasmic reticulum (ER) is a major bottleneck in heterologous protein production. A better understanding of the secretory pathway and the processes involved is therefore of great importance for future optimisation of heterologous protein production in the fungus.

The ER constitutes a major part of the secretory pathway and it is here that protein folding takes place. A quality mechanism ensures that only correctly folded proteins are passed on to the downstream pathway. So when the folding rate of proteins in the ER is lower than the input of newly synthesised polypeptides, unfolded proteins start to accumulate within the lumen of the ER. This constitutes a threat to the cell, and the eukaryotic cell responds to it by the activation of a stress response pathway known as the Unfolded Protein Response (UPR). The UPR consists of a signal transduction cascade that transmits a signal from the lumen of the ER to the nucleus of the cell. The major result of the response is that the cell increases the production of ER localised chaperones and foldases, which are molecules that assist in protein folding.

In this thesis we focused on the properties and the regulation of the transcriptional activator of the *A. niger* UPR, HacA.

Chapter 1 of this thesis gives an overview of the processes and mechanisms within the secretory pathway, and describes the UPR in *Saccharomyces cerevisiae* and mammalian cells. **Chapter 2** describes the cloning and characterisation of the *A. niger hacA* gene. The induced HacA^{i} protein shares 72% identical residues in its DNA binding domain with its yeast homologue, but exhibits only 25% over-all identity. Also the unconventional intron differs substantially. In yeast a 252 nt unconventional intron is removed from the *HAC1*^{u} transcript upon secretion stress, whereas the unconventional intron splicing of *A. niger*

hacA^{u} only removes a 20 nt intron. As a consequence of the unconventional intron splicing, the reading frame downstream of the intron is changed. The reading frame downstream of the intron of *hacA*^{u} encodes a C-terminal domain of 222 amino whereas in *hacA*^{i} a smaller C-terminal domain of 128 amino acids is encoded. In addition to the unconventional intron splicing it was shown that the *A. niger hacA* transcript is truncated by 230 nt upon ER stress. GFP tagging of HacA^{i} proved its nuclear localisation, and over-expression of the spliced (and thus induced) *hacA*^{i} from the *gpdA* promoter resulted in constitutive upregulation of ER localised chaperones and foldases, and induced the transcription of the *hacA* gene itself as well. The *hacA*^{i} gene was also expressed in *E. coli* and protein was purified and used for DNA binding experiments. It was shown by Electrophoretic Mobility Shift Assays (EMSAs) that HacA^{i} was able to bind to promoter fragments of ER localised foldases and chaperones, and to its own promoter.

In **Chapter 3** we focused more on the target sites of the transcriptional activator HacA. The sequence and characteristics of various Unfolded Protein Response Elements (UPREs) was unravelled. The promoters to which HacA^{i} binding was observed in Chapter 1 each contained a single UPRE, except the *hacA* promoter itself in which three UPREs were discovered. A comparison of the identified UPREs resulted in a consensus sequence for the *A. niger* UPRE corresponding to 5'-**CAN(G/A)NTGT/GCCT**-3'. The *A. niger* UPRE is a partially palindromic sequence (underlined) around a single spacer nucleotide (between brackets) followed by a conserved stretch of four nucleotides. By binding site selection (BSS) the optimal binding site for HacA was isolated which shares the characteristics of the natural UPREs, and only differs by the absence of the spacer nucleotide. The dissociation constants for the different UPREs were determined, and by reporter constructs it was shown that UPREs with different dissociation constants resulted in differential upregulation upon ER stress. In addition to the unconventional intron splicing, the *A. niger hacA* transcript is truncated at its 5'end by approximately 230 nt upon secretion stress. The mechanism behind this truncation has been unknown thus far, and has also not been reported for the UPR transcription factors of mammalian cells and yeast. In **Chapter 4** we describe the mechanism behind the *hacA* mRNA truncation.The shorter *hacA* mRNA that appears during secretion stress is the result of transcription of the *hacA* gene from a new start site which lies closer to the ATG. This transcriptional switch is dependent on the UPRE2 in the *hacA* promoter and operated by the HacA^{i} protein itself. The shorter *hacA* transcript lacks the 5'-end half-site of a long inverted repeat that causes translational attenuation in the *hacA*^{u} transcript. Relieved from this translational block, the *hacA*^{i} mRNA can be efficiently translated and contribute to the pool of HacA^{i} protein, thereby strengthen the stress response. Based on our current knowledge a working model is presented at the end of the Chapter, showing the multistep activation of HacA. Finally, in **Chapter 5** the most important results are discussed and a role for HacA^{u} as a negative regulator for the UPR is speculated upon.

Samenvatting

Aspergillus niger is a *filamenteuze* schimmel die wereldwijd verspreid is en veelvuldig voorkomt op rottend planten materiaal. Als een saprophyt speelt *A. niger* een belangrijke rol in de afbraak en recycling van dood plantaardig materiaal. Plantaardig materiaal bestaat voor een groot deel uit complexe polymeren. Om die complexe polymeren te kunnen afbreken tot kleinere moleculen die de schimmel kan opnemen om zo die dienen als nutriënt en energie bron, scheidt *A. niger* een scala aan hydrolytische enzymen uit. Veel van de extracellulaire enzymen van *A. niger*, zoals amylases, proteases, pectinases, xylanases and lipases, vinden toepassingen in de bakkerij, zetmeel, textiel, voedingsmiddelen en diervoeder industrie. In industriële fermentaties kunnen deze extracellulaire enzymen in grote hoeveelheden geproduceerd worden. Voor sommige enzymen worden titers gehaald van meer dan 40g/l. *A. niger* beschikt dus van nature over een secretie pathway met een grote capaciteit, en deze wordt vaak met succes geëxploiteerd voor de productie van homologe enzymen en enzymen van gerelateerde schimmels. De productie van heterologe enzymen van ongerelateerde organismen is vaak echter vele orders van grote lager. Het wordt algemeen aangenomen dat eiwit vouwing in het endoplasmatisch reticulum het belangrijkste obstakel is in heterologe enzym productie met behulp van *A. niger*. Voor de toekomstige optimalisatie van de schimmel voor de productie van heterologe eiwitten, is het daarom van groot belang om een beter inzicht en begrip te krijgen van de secretie pathway en de processen die zich hierin afspelen. Het ER is één van de belangrijkste compartimenten van de secretie pathway, en extra-cellulaire enzymen krijgen hier hun tertiaire structuur door middel van eiwit vouwing. Een kwaliteitscontrole-systeem zorgt ervoor dat alleen correct gevouwen eiwitten worden doorgegeven aan het volgende compartiment in de secretie pathway, het Golgi. Als de snelheid van eiwit vouwing lager is dan de toevoer van nieuw gesynthetiseerde polypeptides, zullen ongevouwen en/of gedeeltelijk gevouwen eiwitten zich dus ophopen in het ER. Dit vormt een serieuze bedreiging voor de cel, en de eukaryote cel reageert hierop door middel van de activering van een stress respons, die bekend staat als de Unfolded Protein Response (UPR). De UPR bestaat uit een signaal transductie cascade dat een signaal stuurt vanuit het ER naar de kern van de cel. Het belangrijkste effect van de signaal overdracht is de verhoogde transcriptie van genen die coderen voor ER gesitueerde chaperones en foldases. Chaperones en foldases zijn moleculen die assisteren bij eiwit vouwing, en het belangrijkste resultaat van de UPR is daardoor dat de capaciteit van eiwit vouwing in de ER wordt verhoogd.

Dit proefschrift is gericht op de ontrafeling van de eigenschappen en de regulering van HacA, de transcriptie factor van de *A. niger* UPR.

Hoofdstuk 1 van dit proefschrift geeft een overzicht van de processen en mechanismen die zich afspelen in de secretie pathway, en beschrijft de UPR zoals die bekend is in de gist *Saccharomyces cerevisiae* and dierlijke cellen.

Hoofdstuk 2 beschrijft de klonering en karakterisatie van het *A. niger hacA* gen. Op eiwit niveau deelt het DNA-bindings domein van HacA^{i} 72% identieke aminozuren met het DNA-bindings domein van zijn gist homoloog, terwijl over de gehele sequentie de identiteit slecht 25% bedraagt. Het onconventionele intron van *hacA* verschil ook substantieel van het intron in het gist *HAC1* mRNA. In gist wordt een 252-nt groot intron uit het *HAC1*^{u} transcript gesplitst gedurende secretie stress, terwijl het *A. niger hacA*^{u} slechts een 20-nt lang intron verliest. Als gevolg van deze onconventionele intron verwijdering, verandert het leesraam van het gedeelte van het transcript volgend op het intron. Terwijl het gedeelte van *hacA*^{u} mRNA na het intron codeert voor een C-terminaal domein van 222 aminozuren, is dit in het intronloze *hacA*^{i} slechts 128 aminozuren. Voor de gist Hac1p is de verandering als gevolg van intron verwijdering minder dramatisch: de laatste 10 aminozuren in Hac1p$^{\underline{u}}$ worden vervangen door fragment van 18 aminozuren in Hac1p^{i}. Naast de onconventionele intron splitsing, werd aangetoond dat het *hacA* transcript 230 nt verliest van zijn 5'-uiteinde, tijdens secretie stress. GFP tagging toonde aan dat HacA^{i} gelokaliseerd is in de kern, en over-expressie van *hacA*^{i} vanaf de constitutieve *gpdA* promotor resulteerde niet alleen in de constitutieve op-regulatie van de ER gebonden chaperones en foldases, maar ook van *hacA* zelf. Het *hacA*^{i} gen werd ook tot expressie gebracht in *E. coli*, en met het vervolgens gezuiverde HacA^{i} eiwit werden DNA bindings experimenten gedaan. Er werd aangetoond dat HacA^{i} in staat was te binden aan promotor fragmenten van ER gelokaliseerde foldases en chaperones, en bovendien aan zijn eigen promotor.

In **Hoofdstuk 3** is vervolgens meer gefocusseerd op de precieze aangrijpingsplaats van HacA in de promotor. De sequenties van deze zogenaamde UPR elementen (UPREs) is opgehelderd, en de bindings-eigenschappen van HacA aan verschillende UPRREs is bepaald. De promotors waarvan in Hoofdstuk 2 was aangetoond dat HacA aan hen kon binden, bleken allen een UPRE te bevatten. De *hacA* promotor zelf bleek zelfs 3 UPREs te bevatten. Een vergelijking van de verschillende geïdentificeerde UPREs leverde de volgende consensus sequentie op: 5'-**<u>CA</u>N(G/A)N<u>TG</u>T/GCCT**-3'

The *A. niger* UPRE is een gedeeltelijk palindromische sequentie (onderstreept) rond een enkele spacer nucleotide (tussen haakjes), gevolgd door een geconserveerde reeks van vier nucleotiden. Door middel van binding-site-selection (BSS), is de sequentie van de optimale bindingsplek van HacA opgehelderd. Deze sequentie verschilt van de natuurlijke UPREs door de afwezigheid van de spacer nucleotide, maar deelt verder de karakteristieken van de UPREs. De dissociatie constante voor de verschillende UPREs is bepaald, en met behulp van reporter constructen is aangetoond dat UPREs met

verschillende dissociatie constanten resulteren in differentiële op-regulatie in geval van secretie stress. Naast de onconventionele intron splitsing, verliest het *A. niger hacA* transscript ook ongeveer 230 nt van zijn 5'-uiteinde gedurende secretie stress. Het mechanisme achter deze verkorting van het mRNA was onbekend tot nog toe, en is ook niet gerapporteerd voor de UPR transscriptie factoren van gist en dierlijke cellen. In **Hoofdstuk 4** wordt het mechanisme achter de *hacA* mRNA verkorting bloot gelegd. De verkorte variant die verschijnt tijdens secretie stress is het gevolg van transcriptie van het *hacA* gen van een nieuwe start plaats die dichter bij het start codon ligt. Deze transcriptionele switch is afhankelijk van de UPRE2 in de *hacA* promotor, en van het HacA^{i} eiwit zelf. Het kortere *hacA^{i}* mRNA mist het 5'-gedeelte van een lange inverted repeat. De interactie van de twee helften van deze inverted repeat zorgt in *hacA^{u}* mRNA voor een translationele blokkade. In *hacA^{i}* is deze blokkade echter opgeheven, en kan het mRNA efficiënt getranslateerd worden om zo bijdragen aan de concentratie van actief HacA^{i} eiwit, en zo de transcriptionele response van de UPR versterken. Op basis van onze huidige kennis is een model gepresenteerd die de multi-staps activatering van HacA beschrijft. Tenslotte zijn in **Hoofdstuk 5** de belangrijkste bevindingen bediscussieerd en wordt een mogelijke rol voorgesteld voor HacA^{u} als negatieve regulator van de *A. niger* UPR.

References

1. **Al Sheikh, H., A. J. Watson, G. A. Lacey, P. J. Punt, D. A. MacKenzie, D. J. Jeenes, T. Pakula, M. Penttila, M. J. Alcocer, and D. B. Archer**. Endoplasmic reticulum stress leads to the selective transcriptional downregulation of the glucoamylase gene in *Aspergillus niger*. *Mol. Microbiol.* 53:1731-1742 (2004).

2. **Archer, D. B. and J. F. Peberdy**. The molecular biology of secreted enzyme production by fungi. *Crit Rev. Biotechnol.* 17:273-306 (1997).

3. **Arnold, S. M., L. I. Fessler, J. H. Fessler, and R. J. Kaufman**. Two homologues encoding human UDP-glucose:glycoprotein glucosyltransferase differ in mRNA expression and enzymatic activity. *Biochemistry* 39:2149-2163 (2000).

4. **Arvas, M., T. Pakula, K. Lanthaler, M. Saloheimo, M. Valkonen, T. Suortti, G. Robson, and M. Penttila**. Common features and interesting differences in transcriptional responses to secretion stress in the fungi *Trichoderma reesei* and *Saccharomyces cerevisiae*. *BMC. Genomics* 7:32 (2006).

5. **Barlowe, C.** Signals for COPII-dependent export from the ER: what's the ticket out? *Trends Cell Biol.* 13:295-300 (2003).

6. **Barlowe, C., L. Orci, T. Yeung, M. Hosobuchi, S. Hamamoto, N. Salama, M. F. Rexach, M. Ravazzola, M. Amherdt, and R. Schekman**. COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77:895-907 (1994).

7. **Belden, W. J. and C. Barlowe**. Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science* 294:1528-1531 (2001).

8. **Bertolotti, A., Y. Zhang, L. M. Hendershot, H. P. Harding, and D. Ron**. Dynamic interaction of BiP and ER stress transducers in the unfolded- protein response. *Nat. Cell Biol.* 2:326-332 (2000).

9. **Bradford, M. M.** A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248-254 (1976).

10. **Cabrera, M., M. Muniz, J. Hidalgo, L. Vega, M. E. Martin, and A. Velasco**. The retrieval function of the KDEL receptor requires PKA phosphorylation of its C-terminus. *Mol. Biol. Cell* 14:4114-4125 (2003).

11. **Caldwell, S. R., K. J. Hill, and A. A. Cooper**. Degradation of endoplasmic reticulum (ER) quality control substrates requires transport between the ER and Golgi. *J. Biol. Chem.* 276:23296-23303 (2001).

12. **Calfon, M., H. Zeng, F. Urano, J. H. Till, S. R. Hubbard, H. P. Harding, S. G. Clark, and D. Ron**. IRE1 couples endoplasmic reticulum load to secretory capacity by processing the *XBP-1* mRNA. *Nature* 415:92-96 (2002).

13. **Chang, S. H., I. Barbosa-Tessmann, C. Chen, M. S. Kilberg, and A. Agarwal**. Glucose deprivation induces heme oxygenase-1 gene expression by a pathway independent of the unfolded protein response. *J. Biol. Chem.* 277:1933-1940 (2002).

14. **Chapman, R. E. and P. Walter**. Translational attenuation mediated by an mRNA intron. *Curr. Biol.* 7:850-859 (1997).

15. **Chen, X., J. Shen, and R. Prywes**. The luminal domain of ATF6 senses endoplasmic reticulum (ER) stress and causes translocation of ATF6 from the ER to the Golgi. *J. Biol. Chem.* 277:13045-13052 (2002).

16. **Clark, M. W. and J. Abelson**. The subnuclear localization of tRNA ligase in yeast. *J. Cell Biol.* 105:1515-1526 (1987).

17. **Conesa, A., D. Jeenes, D. B. Archer, C. A. van den Hondel, and P. J. Punt**. Calnexin overexpression increases manganese peroxidase production in *Aspergillus niger*. *Appl. Environ. Microbiol.* 68:846-851 (2002).

18. **Conesa, A., P. J. Punt, N. van Luijk, and C. A. van den Hondel**. The secretion pathway in filamentous fungi: a biotechnological view. *Fungal. Genet. Biol.* 33:155-171 (2001).

19. **Cosson, P. and F. Letourneur**. Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. *Science* 263:1629-1631 (1994).

20. **Cosson, P. and F. Letourneur**. Coatomer (COPI)-coated vesicles: role in intracellular transport and protein sorting. *Curr. Opin. Cell Biol.* 9:484-487 (1997).

21. **Cox, J. S., C. E. Shamu, and P. Walter**. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* 73:1197-1206 (1993).

22. **Cox, J. S. and P. Walter**. A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. *Cell* 87:391-404 (1996).

23. **Credle, J. J., J. S. Finer-Moore, F. R. Papa, R. M. Stroud, and P. Walter**. On the mechanism of sensing unfolded protein in the endoplasmic reticulum. *Proc. Natl. Acad. Sci U. S. A* 102:18773-18784 (2005).

24. **Cullinan, S. B., D. Zhang, M. Hannink, E. Arvisais, R. J. Kaufman, and J. A. Diehl**. Nrf2 is a direct PERK substrate and effector of PERK-dependent cell survival. *Mol. Cell Biol.* 23:7198-7209 (2003).

25. **de Groot, M. J., P. Bundock, P. J. Hooykaas, and A. G. Beijersbergen**. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol.* 16:839-842 (1998).

26. **Derkx, P. M. and S. M. Madrid**. The *Aspergillus niger cypA* gene encodes a cyclophilin that mediates sensitivity to the immunosuppressant cyclosporin A. *Mol. Genet. Genomics* 266:527-536 (2001).

27. **Derkx, P. M. and S. M. Madrid**. The foldase CYPB is a component of the secretory pathway of *Aspergillus niger* and contains the endoplasmic reticulum retention signal HEEL. *Mol. Genet. Genomics* 266:537-545 (2001).

28. **Ellgaard, L. and A. Helenius**. ER quality control: towards an understanding at the molecular level. *Curr. Opin. Cell Biol.* 13:431-437 (2001).

29. **Ellgaard, L. and A. Helenius**. Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* 4:181-191 (2003).

30. **Ellgaard, L., M. Molinari, and A. Helenius**. Setting the standards: quality control in the secretory pathway. *Science* 286:1882-1888 (1999).

31. **Evans, E. A., R. Gilmore, and G. Blobel**. Purification of microsomal signal peptidase as a complex. *Proc. Natl. Acad. Sci U. S. A* 83:581-585 (1986).

32. **Fassio, A. and R. Sitia**. Formation, isomerisation and reduction of disulphide bonds during protein quality control in the endoplasmic reticulum. *Histochem. Cell Biol.* 117:151-157 (2002).

33. **Fornerod, M., M. Ohno, M. Yoshida, and I. W. Mattaj**. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90:1051-1060 (1997).

34. **Forsee, W. T., J. D. Springfield, and J. S. Schutzbach**. Effect of phospholipids on alpha-1,2-mannosidase activity. *J. Biol. Chem.* 257:9963-9967 (1982).

35. **Foulds, G. J. and F. A. Etzkorn**. A capillary electrophoresis mobility shift assay for protein-DNA binding affinities free in solution. *Nucleic Acids Res.* 26:4304-4305 (1998).

36. **Frand, A. R. and C. A. Kaiser**. The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Mol. Cell* 1:161-170 (1998).

37. **Galagan, J. E., S. E. Calvo, C. Cuomo, L. J. Ma, J. R. Wortman, S. Batzoglou, S. I. Lee, M. Basturkmen, C. C. Spevak, J. Clutterbuck, V. Kapitonov, J. Jurka, C. Scazzocchio, M. Farman, J. Butler, S. Purcell, S. Harris, G. H. Braus, O. Draht, S. Busch, C. D'Enfert, C. Bouchier, G. H. Goldman, D. Bell-Pedersen, S. Griffiths-Jones, J. H. Doonan, J. Yu, K. Vienken, A. Pain, M. Freitag, E. U. Selker, D. B. Archer, M. A. Penalva, B. R. Oakley, M. Momany, T. Tanaka, T. Kumagai, K. Asai, M. Machida, W. C. Nierman, D. W. Denning, M. Caddick, M. Hynes, M. Paoletti, R. Fischer, B. Miller, P. Dyer, M. S. Sachs, S. A. Osmani, and B. W. Birren**. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438:1105-1115 (2005).

38. **Gasch, A. P., P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D. Botstein, and P. O. Brown**. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* 11:4241-4257 (2000).

39. **Gething, M. J. and J. Sambrook**. Protein folding in the cell. *Nature* 355:33-45 (1992).

40. **Geysens, S., G. Whyteside, and D. Archer**. Genomics of protein folding in the endoplasmic reticulum, secretion stress and glycosylation in the aspergilli. *Fungal. Genet. Biol.* 46 Suppl 1:S121-S140 (2009).

41. **Glick, B. S.** Can Hsp70 proteins act as force-generating motors? *Cell* 80:11-14 (1995).

42. **Goffin, L., S. Vodala, C. Fraser, J. Ryan, M. Timms, S. Meusburger, B. Catimel, E. C. Nice, P. A. Silver, C. Y. Xiao, D. A. Jans, and M. J. Gething**. The unfolded protein response transducer Ire1p contains a nuclear localization sequence recognized by multiple beta importins. *Mol. Biol. Cell* 17:5309-5323 (2006).

43. **Gonzalez, T. N., C. Sidrauski, S. Dorfler, and P. Walter**. Mechanism of non-spliceosomal mRNA splicing in the unfolded protein response pathway. *EMBO J.* 18:3119-3132 (1999).

44. **Goodbourn, S.** 1996. Eukaryotic gene transcription. Oxford University Press, Oxford.

45. **Goosen, T., G. Bloemheuvel, C. Gysler, D. A. de Bie, H. W. van den Broek, and K. Swart**. Transformation of *Aspergillus niger* using the homologous orotidine-5'-phosphate-decarboxylase gene. *Curr. Genet.* 11:499-503 (1987).

46. **Gordon, C. L., D. B. Archer, D. J. Jeenes, J. H. Doonan, B. Wells, A. P. Trinci, and G. D. Robson**. A glucoamylase::GFP gene fusion to study protein secretion by individual hyphae of *Aspergillus niger*. *J. Microbiol. Methods* 42:39-48 (2000).

47. **Gouka, R. J., P. J. Punt, and C. A. van den Hondel**. Efficient production of secreted proteins by *Aspergillus*: progress, limitations and prospects. *Appl. Microbiol. Biotechnol.* 47:1-11 (1997).

48. **Guillemette, T., N. N. van Peij, T. Goosen, K. Lanthaler, G. D. Robson, C. A. van den Hondel, H. Stam, and D. B. Archer**. Genomic analysis of the secretion stress response in the enzyme-producing cell factory *Aspergillus niger*. *BMC. Genomics* 8:158 (2007).

49. **Hammond, C., I. Braakman, and A. Helenius**. Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proc. Natl. Acad. Sci U. S. A* 91:913-917 (1994).

50. **Harding, H. P., M. Calfon, F. Urano, I. Novoa, and D. Ron**. Transcriptional and translational control in the Mammalian unfolded protein response. *Annu. Rev. Cell Dev. Biol.* 18:575-599 (2002).

51. **Harding, H. P., I. Novoa, Y. Zhang, H. Zeng, R. Wek, M. Schapira, and D. Ron**. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Mol. Cell* 6:1099-1108 (2000).

52. **Harding, H. P., Y. Zhang, and D. Ron**. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* 397:271-274 (1999).

53. **Harding, H. P., Y. Zhang, H. Zeng, I. Novoa, P. D. Lu, M. Calfon, N. Sadri, C. Yun, B. Popko, R. Paules, D. F. Stojdl, J. C. Bell, T. Hettmann, J. M. Leiden, and D. Ron**. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol. Cell* 11:619-633 (2003).

54. **Haze, K., H. Yoshida, H. Yanagi, T. Yura, and K. Mori**. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol. Biol. Cell* 10:3787-3799 (1999).

55. **Helenius, A. and M. Aebi**. Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* 73:1019-1049 (2004).

56. **Helenius, A., T. Marquardt, and I. Braakman**. The endoplasmic reticulum as a protein-folding compartment. *Trends Cell Biol.* 2:227-231 (1992).

57. **High, S., F. J. Lecomte, S. J. Russell, B. M. Abell, and J. D. Oliver**. Glycoprotein folding in the endoplasmic reticulum: a tale of three chaperones? *FEBS Lett.* 476:38-41 (2000).

58. **Hijarrubia, M. J., J. Casqueiro, S. Gutierrez, F. J. Fernandez, and J. F. Martin**. Characterization of the *bip* gene of *Aspergillus awamori* encoding a protein with an HDEL retention signal homologous to the mammalian BiP involved in polypeptide secretion. *Curr. Genet.* 32:139-146 (1997).

59. **Hinnebusch, A. G.** Translational regulation of yeast *GCN4*. A window on factors that control initiator-tRNA binding to the ribosome. *J. Biol. Chem.* 272:21661-21664 (1997).

60. **Isler, J. A., A. H. Skalet, and J. C. Alwine**. Human cytomegalovirus infection activates and regulates the unfolded protein response. *J. Virol.* 79:6890-6899 (2005).

61. **Itoh, K., N. Wakabayashi, Y. Katoh, T. Ishii, K. Igarashi, J. D. Engel, and M. Yamamoto**. Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev.* 13:76-86 (1999).

62. **Jeenes, D. J., R. Pfaller, and D. B. Archer**. Isolation and characterisation of a novel stress-inducible PDI-family gene from *Aspergillus niger*. *Gene* 193:151-156 (1997).

63. **Jorgensen, T. R., T. Goosen, C. A. Hondel, A. F. Ram, and J. J. Iversen**. Transcriptomic comparison of *Aspergillus niger* growing on two different sugars reveals coordinated regulation of the secretory pathway. *BMC. Genomics* 10:44 (2009).

64. **Kaufman, R. J.** Stress signaling from the lumen of the endoplasmic reticulum: coordination of gene transcriptional and translational controls. *Genes Dev.* 13:1211-1233 (1999).

65. **Kaufman, R. J.** Regulation of mRNA translation by protein folding in the endoplasmic reticulum. *Trends Biochem. Sci.* 29:152-158 (2004).

66. **Kawahara, T., H. Yanagi, T. Yura, and K. Mori**. Endoplasmic reticulum stress-induced mRNA splicing permits synthesis of transcription factor Hac1p/Ern4p that activates the unfolded protein response. *Mol. Biol. Cell* 8:1845-1862 (1997).

67. **Kawahara, T., H. Yanagi, T. Yura, and K. Mori**. Unconventional splicing of *HAC1/ERN4* mRNA required for the unfolded protein response. Sequence-specific and non-sequential cleavage of the splice sites. *J. Biol. Chem.* 273:1802-1807 (1998).

68. **Kent, C.** Eukaryotic phospholipid biosynthesis. *Annu. Rev. Biochem.* 64:315-343 (1995).

69. **Kim, T. Y., E. Kim, S. K. Yoon, and J. B. Yoon**. Herp enhances ER-associated protein degradation by recruiting ubiquilins. *Biochem. Biophys. Res. Commun.* 369:741-746 (2008).

70. **Kimata, Y., Y. I. Kimata, Y. Shimizu, H. Abe, I. C. Farcasanu, M. Takeuchi, M. D. Rose, and K. Kohno**. Genetic evidence for a role of BiP/Kar2 that regulates Ire1 in response to accumulation of unfolded proteins. *Mol. Biol. Cell* 14:2559-2569 (2003).

71. **Kimata, Y., D. Oikawa, Y. Shimizu, Y. Ishiwata-Kimata, and K. Kohno**. A role for BiP as an adjustor for the endoplasmic reticulum stress-sensing protein Ire1. *J. Cell Biol.* 167:445-456 (2004).

72. **Kohno, K., K. Normington, J. Sambrook, M. J. Gething, and K. Mori**. The promoter region of the yeast *KAR2* (BiP) gene contains a regulatory domain that responds to the presence of unfolded proteins in the endoplasmic reticulum. *Mol. Cell Biol.* 13:877-890 (1993).

73. **Kokame, K., K. L. Agarwala, H. Kato, and T. Miyata**. Herp, a new ubiquitin-like membrane protein induced by endoplasmic reticulum stress. *J. Biol. Chem.* 275:32846-32853 (2000).

74. **Kokame, K., H. Kato, and T. Miyata**. Identification of ERSE-II, a new *cis*-acting element responsible for the ATF6-dependent mammalian unfolded protein response. *J. Biol. Chem.* 276:9199-9205 (2001).

75. **Kruse, K. B., J. L. Brodsky, and A. A. McCracken**. Autophagy: an ER protein quality control process. *Autophagy.* 2:135-137 (2006).

76. **Kusters-van Someren, M. A., J. A. Harmsen, H. C. Kester, and J. Visser**. Structure of the *Aspergillus niger pel*A gene and its expression in *Aspergillus niger* and *Aspergillus nidulans*. *Curr. Genet.* 20:293-299 (1991).

77. **la Cour, T., L. Kiemer, A. Molgaard, R. Gupta, K. Skriver, and S. Brunak**. Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel* 17:527-536 (2004).

78. **Latchman, D.** 1995. Gene regulation: A eukaryotic perspective. Chapman & Hall, London.

79. **Leber, J. H., S. Bernales, and P. Walter**. IRE1-independent gain control of the unfolded protein response. *PLoS. Biol.* 2:E235 (2004).

80. **Lee, A. H., N. N. Iwakoshi, and L. H. Glimcher**. XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. *Mol. Cell Biol.* 23:7448-7459 (2003).

81. **Lee, K., W. Tirasophon, X. Shen, M. Michalak, R. Prywes, T. Okada, H. Yoshida, K. Mori, and R. J. Kaufman**. IRE1-mediated unconventional mRNA splicing and S2P-mediated ATF6 cleavage merge to regulate XBP1 in signaling the unfolded protein response. *Genes Dev.* 16:452-466 (2002).

82. **Lee, M. C., E. A. Miller, J. Goldberg, L. Orci, and R. Schekman**. Bi-directional protein transport between the ER and Golgi. *Annu. Rev. Cell Dev. Biol.* 20:87-123 (2004).

83. **Lewis, M. J. and H. R. Pelham**. Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. *Cell* 68:353-364 (1992).

84. **Liu, C. Y., M. Schroder, and R. J. Kaufman**. Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum. *J. Biol. Chem.* 275:24881-24885 (2000).

85. **Ma, Y. and L. M. Hendershot**. The role of the unfolded protein response in tumour development: friend or foe? *Nat. Rev. Cancer* 4:966-977 (2004).

86. **MacArthur, M. W. and J. M. Thornton**. Influence of proline residues on protein conformation. *J. Mol. Biol.* 218:397-412 (1991).

87. **Machida, M., K. Asai, M. Sano, T. Tanaka, T. Kumagai, G. Terai, K. Kusumoto, T. Arima, O. Akita, Y. Kashiwagi, K. Abe, K. Gomi, H. Horiuchi, K. Kitamoto, T. Kobayashi, M. Takeuchi, D. W. Denning, J. E. Galagan, W. C. Nierman, J. Yu, D. B. Archer, J. W. Bennett, D. Bhatnagar, T. E. Cleveland, N. D. Fedorova, O. Gotoh, H. Horikawa, A. Hosoyama, M. Ichinomiya, R. Igarashi, K. Iwashita, P. R. Juvvadi, M. Kato, Y. Kato, T. Kin, A. Kokubun, H. Maeda, N. Maeyama, J. Maruyama, H. Nagasaki, T. Nakajima, K. Oda, K. Okada, I. Paulsen, K. Sakamoto, T. Sawano, M. Takahashi, K. Takase, Y. Terabayashi, J. R. Wortman, O. Yamada, Y. Yamagata, H. Anazawa, Y. Hata, Y. Koide, T. Komori, Y. Koyama, T. Minetoki, S. Suharnan, A. Tanaka, K. Isono, S. Kuhara, N. Ogasawara, and H. Kikuchi**. Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 438:1157-1161 (2005).

88. **Maras, M., A. De Bruyn, J. Schraml, P. Herdewijn, M. Claeyssens, W. Fiers, and R. Contreras**. Structural characterization of N-linked oligosaccharides from cellobiohydrolase I secreted by the filamentous fungus *Trichoderma reesei* RUTC 30. *Eur. J. Biochem.* 245:617-625 (1997).

89. **Martinez, I. M. and M. J. Chrispeels**. Genomic analysis of the unfolded protein response in *Arabidopsis* shows its connection to important cellular processes. *Plant Cell* 15:561-576 (2003).

90. **Meijer, H. A. and A. A. Thomas**. Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem. J.* 367:1-11 (2002).

91. **Meusser, B., C. Hirsch, E. Jarosch, and T. Sommer**. ERAD: the long road to destruction. *Nat. Cell Biol.* 7:766-772 (2005).

92. **Michielse, C. B., M. Arentshorst, A. F. Ram, and C. A. van den Hondel**. *Agrobacterium*-mediated transformation leads to improved gene replacement efficiency in *Aspergillus awamori*. *Fungal. Genet. Biol.* 42:9-19 (2005).

93. **Miyazaki, M. and J. M. Ntambi**. Role of stearoyl-coenzyme A desaturase in lipid metabolism. *Prostaglandins Leukot. Essent. Fatty Acids* 68:113-121 (2003).

94. **Mori, K.** Tripartite management of unfolded proteins in the endoplasmic reticulum. *Cell* 101:451-454 (2000).

95. **Mori, K., T. Kawahara, H. Yoshida, H. Yanagi, and T. Yura**. Signalling from endoplasmic reticulum to nucleus: transcription factor with a basic-leucine zipper motif is required for the unfolded protein- response pathway. *Genes Cells* 1:803-817 (1996).

96. **Mori, K., W. Ma, M. J. Gething, and J. Sambrook**. A transmembrane protein with a cdc2+/CDC28-related kinase activity is required for signaling from the ER to the nucleus. *Cell* 74:743-756 (1993).

97. **Mori, K., N. Ogawa, T. Kawahara, H. Yanagi, and T. Yura**. Palindrome with spacer of one nucleotide is characteristic of the *cis*-acting unfolded protein response element in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 273:9912-9920 (1998).

98. **Mori, K., N. Ogawa, T. Kawahara, H. Yanagi, and T. Yura**. mRNA splicing-mediated C-terminal replacement of transcription factor Hac1p is required for efficient activation of the unfolded protein response. *Proc. Natl. Acad. Sci. U. S. A* 97:4660-4665 (2000).

99. **Mori, K., A. Sant, K. Kohno, K. Normington, M. J. Gething, and J. F. Sambrook**. A 22 bp *cis*-acting element is necessary and sufficient for the induction of the yeast *KAR2* (BiP) gene by unfolded proteins. *EMBO J.* 11:2583-2593 (1992).

100. **Mulder, H. J., I. Nikolaev, and S. M. Madrid**. HACA, the transcriptional activator of the unfolded protein response (UPR) in *Aspergillus niger*, binds to partly palindromic UPR elements of the consensus sequence 5'-CAN(G/A)NTGT/GCCT-3'. *Fungal. Genet. Biol.* 43:560-572 (2006).

101. **Mulder, H. J., M. Saloheimo, M. Penttila, and S. M. Madrid**. The transcription factor HACA mediates the unfolded protein response in *Aspergillus niger*, and up-regulates its own transcription. *Mol. Genet. Genomics* 271:130-140 (2004).

102. **Natarajan, K., M. R. Meyer, B. M. Jackson, D. Slade, C. Roberts, A. G. Hinnebusch, and M. J. Marton**. Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. *Mol. Cell Biol.* 21:4347-4368 (2001).

103. **Ng, D. T., J. D. Brown, and P. Walter**. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J. Cell Biol.* 134:269-278 (1996).

104. **Ngiam, C., D. J. Jeenes, and D. B. Archer**. Isolation and characterisation of a gene encoding protein disulphide isomerase, *pdiA*, from *Aspergillus niger*. *Curr. Genet.* 31:133-138 (1997).

105. **Ngiam, C., D. J. Jeenes, P. J. Punt, C. A. van den Hondel, and D. B. Archer**. Characterization of a foldase, protein disulfide isomerase A, in the protein secretory pathway of *Aspergillus niger*. *Appl. Environ. Microbiol.* 66:775-782 (2000).

106. **Nikawa, J. and S. Yamashita**. *IRE1* encodes a putative protein kinase containing a membrane-spanning domain and is required for inositol phototrophy in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 6:1441-1446 (1992).

107. **Niwa, M., C. K. Patil, J. DeRisi, and P. Walter**. Genome-scale approaches for discovering novel nonconventional splicing substrates of the Ire1 nuclease. *Genome Biol.* 6:R3 (2005).

108. **Niwa, M., C. Sidrauski, R. J. Kaufman, and P. Walter**. A role for presenilin-1 in nuclear accumulation of Ire1 fragments and induction of the mammalian unfolded protein response. *Cell* 99:691-702 (1999).

109. **Noiva, R.** Enzymatic catalysis of disulfide formation. *Protein Expr. Purif.* 5:1-13 (1994).

110. **Nojima, H., S. H. Leem, H. Araki, A. Sakai, N. Nakashima, Y. Kanaoka, and Y. Ono**. Hac1: a novel yeast bZIP protein binding to the CRE motif is a multicopy suppressor for cdc10 mutant of *Schizosaccharomyces pombe*. *Nucleic Acids Res.* 22:5279-5288 (1994).

111. **Offringa, R., M. J. de Groot, H. J. Haagsman, M. P. Does, P. J. van den Elzen, and P. J. Hooykaas**. Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *EMBO J.* 9:3077-3084 (1990).

112. **Ogawa, N. and K. Mori**. Autoregulation of the HAC1 gene is required for sustained activation of the yeast unfolded protein response. *Genes Cells* 9:95-104 (2004).

113. **Oikawa, D., Y. Kimata, and K. Kohno**. Self-association and BiP dissociation are not sufficient for activation of the ER stress sensor Ire1. *J. Cell Sci* 120:1681-1688 (2007).

114. **Oka, T. and A. Nakano**. Inhibition of GTP hydrolysis by Sar1p causes accumulation of vesicles that are a functional intermediate of the ER-to-Golgi transport in yeast. *J. Cell Biol.* 124:425-434 (1994).

115. **Okamura, K., Y. Kimata, H. Higashio, A. Tsuru, and K. Kohno**. Dissociation of Kar2p/BiP from an ER sensory molecule, Ire1p, triggers the unfolded protein response in yeast. *Biochem. Biophys. Res. Commun.* 279:445-450 (2000).

116. **Otsu, M. and R. Sitia**. Diseases originating from altered protein quality control in the endoplasmic reticulum. *Curr. Med. Chem.* 14:1639-1652 (2007).

117. **Ou, W. J., J. J. Bergeron, Y. Li, C. Y. Kang, and D. Y. Thomas**. Conformational changes induced in the endoplasmic reticulum luminal domain of calnexin by Mg-ATP and Ca2+. *J. Biol. Chem.* 270:18051-18059 (1995).

118. **Pakula, T. M., M. Laxell, A. Huuskonen, J. Uusitalo, M. Saloheimo, and M. Penttila**. The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei.* Evidence for down-regulation of genes that encode secreted proteins in the stressed cells. *J. Biol. Chem.* 278:45011-45020 (2003).

119. **Pakula, T. M., M. Laxell, A. Huuskonen, J. Uusitalo, M. Saloheimo, and M. Penttila**. The effects of drugs inhibiting protein secretion in the filamentous fungus *Trichoderma reesei*.

Evidence for down-regulation of genes that encode secreted proteins in the stressed cells. *J. Biol. Chem.* 278:45011-45020 (2003).

120. **Pakula, T. M., J. Uusitalo, M. Saloheimo, K. Salonen, R. J. Aarts, and M. Penttila**. Monitoring the kinetics of glycoprotein synthesis and secretion in the filamentous fungus *Trichoderma reesei*: cellobiohydrolase I (CBHI) as a model protein. *Microbiology* 146 (Pt 1):223-232 (2000).

121. **Parodi, A. J.** Protein glucosylation and its role in protein folding. *Annu. Rev. Biochem.* 69:69-93 (2000).

122. **Partaledis, J. A. and V. Berlin**. The *FKB2* gene of *Saccharomyces cerevisiae*, encoding the immunosuppressant-binding protein FKBP-13, is regulated in response to accumulation of unfolded proteins in the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U. S. A* 90:5450-5454 (1993).

123. **Patil, C. and P. Walter**. Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. *Curr. Opin. Cell Biol.* 13:349-355 (2001).

124. **Patil, C. K., H. Li, and P. Walter**. Gcn4p and novel upstream activating sequences regulate targets of the unfolded protein response. *PLoS. Biol.* 2:E246 (2004).

125. **Peberdy, J. F.** Protein secretion in filamentous fungi--trying to understand a highly productive black box. *Trends Biotechnol.* 12:50-57 (1994).

126. **Pel, H. J., J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann, P. J. Schaap, G. Turner, R. P. de Vries, R. Albang, K. Albermann, M. R. Andersen, J. D. Bendtsen, J. A. Benen, B. M. van den, S. Breestraat, M. X. Caddick, R. Contreras, M. Cornell, P. M. Coutinho, E. G. Danchin, A. J. Debets, P. Dekker, P. W. van Dijck, A. van Dijk, L. Dijkhuizen, A. J. Driessen, C. D'Enfert, S. Geysens, C. Goosen, G. S. Groot, P. W. de Groot, T. Guillemette, B. Henrissat, M. Herweijer, J. P. van den Hombergh, C. A. van den Hondel, R. T. van der Heijden, R. M. van der Kaaij, F. M. Klis, H. J. Kools, C. P. Kubicek, P. A. van Kuyk, J. Lauber, X. Lu, M. J. van der Maarel, R. Meulenberg, H. Menke, M. A. Mortimer, J. Nielsen, S. G. Oliver, M. Olsthoorn, K. Pal, N. N. van Peij, A. F. Ram, U. Rinas, J. A. Roubos, C. M. Sagt, M. Schmoll, J. Sun, D. Ussery, J. Varga, W. Vervecken, P. J. van de Vondervoort, H. Wedler, H. A. Wosten, A. P. Zeng, A. J. van Ooyen, J. Visser, and H. Stam**. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 25:221-231 (2007).

127. **Pelham, H. R.** The dynamic organisation of the secretory pathway. *Cell Struct. Funct.* 21:413-419 (1996).

128. **Pierrou, S., S. Enerback, and P. Carlsson**. Selection of high-affinity binding sites for sequence-specific, DNA binding proteins from random sequence oligonucleotides. *Anal. Biochem.* 229:99-105 (1995).

129. **Punt, P. J., N. van Biezen, A. Conesa, A. Albers, J. Mangnus, and H. C. van den**. Filamentous fungi as cell factories for heterologous protein production. *Trends Biotechnol.* 20:200-206 (2002).

130. **Punt, P. J., I. A. van Gemeren, J. Drint-Kuijvenhoven, J. G. Hessing, G. M. Muijlwijk-Harteveld, A. Beijersbergen, C. T. Verrips, and C. A. van den Hondel**. Analysis of the role of the gene *bipA*, encoding the major endoplasmic reticulum chaperone protein in the

secretion of homologous and heterologous proteins in black *Aspergilli*. *Appl. Microbiol. Biotechnol.* 50:447-454 (1998).

131. **Rapoport, T. A., M. M. Rolls, and B. Jungnickel**. Approaching the mechanism of protein transport across the ER membrane. *Curr. Opin. Cell Biol.* 8:499-504 (1996).

132. **Richie, D. L., L. Hartl, V. Aimanianda, M. S. Winters, K. K. Fuller, M. D. Miley, S. White, J. W. McCarthy, J. P. Latge, M. Feldmesser, J. C. Rhodes, and D. S. Askew**. A role for the unfolded protein response (UPR) in virulence and antifungal susceptibility in *Aspergillus fumigatus*. *PLoS. Pathog.* 5:e1000258 (2009).

133. **Riggs, A. D., H. Suzuki, and S. Bourgeoss**. Lac repressor-operator interaction. I. Equilibrium studies. *J. Mol. Biol.* 48:67-83 (1970).

134. **Ritter, C. and A. Helenius**. Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. *Nat. Struct. Biol.* 7:278-280 (2000).

135. **Romisch, K.** Endoplasmic reticulum-associated degradation. *Annu. Rev. Cell Dev. Biol.* 21:435-456 (2005).

136. **Roy, B. and A. S. Lee**. The mammalian endoplasmic reticulum stress response element consists of an evolutionarily conserved tripartite structure and interacts with a novel stress-inducible complex. *Nucleic Acids Res.* 27:1437-1443 (1999).

137. **Ruegsegger, U., J. H. Leber, and P. Walter**. Block of HAC1 mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response. *Cell* 107:103-114 (2001).

138. **Rutkowski, D. T. and R. J. Kaufman**. A trip to the ER: coping with stress. *Trends Cell Biol.* 14:20-28 (2004).

139. **Saloheimo, M., M. Lund, and M. E. Penttila**. The protein disulphide isomerase gene of the fungus *Trichoderma reesei* is induced by endoplasmic reticulum stress and regulated by the carbon source. *Mol. Gen. Genet.* 262:35-45 (1999).

140. **Saloheimo, M., M. Valkonen, and M. Penttila**. Activation mechanisms of the HACI-mediated unfolded protein response in filamentous fungi. *Mol. Microbiol.* 47:1149-1161 (2003).

141. **Sambrook, J., E. F. Fritsch, and T. Maniatis**. 1989. Molecular cloning: A laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor NY..

142. **Schekman, R. and L. Orci**. Coat proteins and vesicle budding. *Science* 271:1526-1533 (1996).

143. **Scheuner, D., B. Song, E. McEwen, C. Liu, R. Laybutt, P. Gillespie, T. Saunders, S. Bonner-Weir, and R. J. Kaufman**. Translational control is required for the unfolded protein response and *in vivo* glucose homeostasis. *Mol. Cell* 7:1165-1176 (2001).

144. **Schroder, M. and R. J. Kaufman**. The mammalian unfolded protein response. *Annu. Rev. Biochem.* 74:739-789 (2005).

145. **Schuler, G. D., S. F. Altschul, and D. J. Lipman**. A workbench for multiple alignment construction and analysis. *Proteins* 9:180-190 (1991).

146. **Schuster, E., N. Dunn-Coleman, J. C. Frisvad, and P. W. van Dijck**. On the safety of *Aspergillus niger*--a review. *Appl Microbiol. Biotechnol.* 59:426-435 (2002).

147. **Shamu, C. E. and P. Walter**. Oligomerization and phosphorylation of the Ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. *EMBO J.* 15:3028-3039 (1996).

148. **Shen, X., R. E. Ellis, K. Lee, C. Y. Liu, K. Yang, A. Solomon, H. Yoshida, R. Morimoto, D. M. Kurnit, K. Mori, and R. J. Kaufman**. Complementary signaling pathways regulate the unfolded protein response and are required for *C. elegans* development. *Cell* 107:893-903 (2001).

149. **Shoji, J. Y., M. Arioka, and K. Kitamoto**. Dissecting cellular components of the secretory pathway in filamentous fungi: insights into their application for protein production. *Biotechnol. Lett.* 30:7-14 (2008).

150. **Sidrauski, C., R. Chapman, and P. Walter**. The unfolded protein response: an intracellular signalling pathway with many surprising features. *Trends Cell Biol.* 8:245-249 (1998).

151. **Sidrauski, C., J. S. Cox, and P. Walter**. tRNA ligase is required for regulated mRNA splicing in the unfolded protein response. *Cell* 87:405-413 (1996).

152. **Sidrauski, C. and P. Walter**. The transmembrane kinase Ire1p is a site-specific endonuclease that initiates mRNA splicing in the unfolded protein response. *Cell* 90:1031-1039 (1997).

153. **Siebert, P. D., A. Chenchik, D. E. Kellogg, K. A. Lukyanov, and S. A. Lukyanov**. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 23:1087-1088 (1995).

154. **Sims, A. H., M. E. Gent, K. Lanthaler, N. S. Dunn-Coleman, S. G. Oliver, and G. D. Robson**. Transcriptome analysis of recombinant protein secretion by *Aspergillus nidulans* and the unfolded-protein response *in vivo*. *Appl. Environ. Microbiol.* 71:2737-2747 (2005).

155. **Steel, G. J., D. M. Fullerton, J. R. Tyson, and C. J. Stirling**. Coordinated activation of Hsp70 chaperones. *Science* 303:98-101 (2004).

156. **Suzuki, C. K., J. S. Bonifacino, A. Y. Lin, M. M. Davis, and R. D. Klausner**. Regulating the retention of T-cell receptor alpha chain variants within the endoplasmic reticulum: Ca(2+)-dependent association with BiP. *J. Cell Biol.* 114:189-205 (1991).

157. **Takatsuki, A., K. Arima, and G. Tamura**. Tunicamycin, a new antibiotic. I. Isolation and characterization of tunicamycin. *J. Antibiot. (Tokyo)* 24:215-223 (1971).

158. **Tehlivets, O., K. Scheuringer, and S. D. Kohlwein**. Fatty acid synthesis and elongation in yeast. *Biochim. Biophys. Acta* 1771:255-270 (2007).

159. **Tirasophon, W., A. A. Welihinda, and R. J. Kaufman**. A stress response pathway from the endoplasmic reticulum to the nucleus requires a novel bifunctional protein kinase/endoribonuclease (Ire1p) in mammalian cells. *Genes Dev.* 12:1812-1824 (1998).

160. **Travers, K. J., C. K. Patil, L. Wodicka, D. J. Lockhart, J. S. Weissman, and P. Walter**. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* 101:249-258 (2000).

161. **Trusina, A., F. R. Papa, and C. Tang**. Rationalizing translation attenuation in the network architecture of the unfolded protein response. *Proc. Natl. Acad. Sci U. S. A* 105:20280-20285 (2008).

162. **Tu, B. P. and J. S. Weissman**. Oxidative protein folding in eukaryotes: mechanisms and consequences. *J. Cell Biol.* 164:341-346 (2004).

163. **Tyson, J. R. and C. J. Stirling**. LHS1 and SIL1 provide a lumenal function that is essential for protein translocation into the endoplasmic reticulum. *EMBO J.* 19:6440-6452 (2000).

164. **Valkonen, M., M. Penttila, and M. Saloheimo**. Effects of inactivation and constitutive expression of the unfolded- protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Appl Environ. Microbiol.* 69:2065-2072 (2003).

165. **Valkonen, M., M. Penttila, and M. Saloheimo**. The *ire1* and *ptc2* genes involved in the unfolded protein response pathway in the filamentous fungus *Trichoderma reesei*. *Mol. Genet. Genomics* 272:443-451 (2004).

166. **Valkonen, M., M. Ward, H. Wang, M. Penttila, and M. Saloheimo**. Improvement of foreign-protein production in *Aspergillus niger var. awamori* by constitutive induction of the unfolded-protein response. *Appl. Environ. Microbiol.* 69:6979-6986 (2003).

167. **van den Hombergh, J. P., P. J. van de Vondervoort, L. Fraissinet-Tachet, and J. Visser**. *Aspergillus* as a host for heterologous protein production: the problem of proteases. *Trends Biotechnol.* 15:256-263 (1997).

168. **van Gemeren, I. A., P. J. Punt, A. Drint-Kuyvenhoven, M. P. Broekhuijsen, A. van't Hoog, A. Beijersbergen, C. T. Verrips, and C. A. van den Hondel**. The ER chaperone encoding *bipA* gene of black *Aspergilli* is induced by heat shock and unfolded proteins. *Gene* 198:43-52 (1997).

169. **van Hartingsveldt, W., I. E. Mattern, C. M. van Zeijl, P. H. Pouwels, and C. A. van den Hondel**. Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene. *Mol. Gen. Genet.* 206:71-75 (1987).

170. **Verdoes, J. C., P. J. Punt, and C. A. van den Hondel**. Molecular genetic strain improvement for the overproduction of fungal proteins by Filamentous fungi. *Appl Microbiol. Biotechnol.* 43:195-205 (1995).

171. **Vilela, C. and J. E. McCarthy**. Regulation of fungal gene expression via short open reading frames in the mRNA 5'untranslated region. *Mol. Microbiol.* 49:859-867 (2003).

172. **Vishniac, W. and M. Santer**. The *thiobacilli*. *Bacteriol. Rev* 21:195-213 (1957).

173. **Wang, H., J. Entwistle, E. Morlon, D. B. Archer, J. F. Peberdy, M. Ward, and D. J. Jeenes**. Isolation and characterisation of a calnexin homologue, *clxA*, from *Aspergillus niger*. *Mol. Genet. Genomics* 268:684-691 (2003).

174. **Wang, H. and M. Ward**. Molecular characterization of a PDI-related gene *prpA* in *Aspergillus niger* var. *awamori*. *Curr. Genet.* 37:57-64 (2000).

175. **Wang, X. Z., H. P. Harding, Y. Zhang, E. M. Jolicoeur, M. Kuroda, and D. Ron**. Cloning of mammalian Ire1 reveals diversity in the ER stress responses. *EMBO J.* 17:5708-5717 (1998).

176. **Wang, Y., J. Shen, N. Arenzana, W. Tirasophon, R. J. Kaufman, and R. Prywes**. Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. *J. Biol. Chem.* 275:27013-27020 (2000).

177. **Wanke, C., S. Eckert, G. Albrecht, W. van Hartingsveldt, P. J. Punt, C. A. van den Hondel, and G. H. Braus**. The *Aspergillus niger GCN4* homologue, *cpcA*, is transcriptionally regulated and encodes an unusual leucine zipper. *Mol. Microbiol.* 23:23-33 (1997).

178. **Welihinda, A. A. and R. J. Kaufman**. The unfolded protein response pathway in *Saccharomyces cerevisiae*. Oligomerization and trans-phosphorylation of Ire1p (Ern1p) are required for kinase activation. *J. Biol. Chem.* 271:18181-18187 (1996).

179. **Welihinda, A. A., W. Tirasophon, and R. J. Kaufman**. The cellular response to protein misfolding in the endoplasmic reticulum. *Gene Expr.* 7:293-300 (1999).

180. **Welihinda, A. A., W. Tirasophon, and R. J. Kaufman**. The transcriptional co-activator ADA5 is required for HAC1 mRNA processing *in vivo*. *J. Biol. Chem.* 275:3377-3381 (2000).

181. **Wiebe, M. G., A. Karandikar, G. D. Robson, A. P. Trinci, J. L. Candia, S. Trappe, G. Wallis, U. Rinas, P. M. Derkx, S. M. Madrid, H. Sisniega, I. Faus, R. Montijn, C. A. van den Hondel, and P. J. Punt**. Production of tissue plasminogen activator (t-PA) in *Aspergillus niger*. *Biotechnol. Bioeng.* 76:164-174 (2001).

182. **Wosten, H. A., S. M. Moukha, J. H. Sietsma, and J. G. Wessels**. Localization of growth and secretion of proteins in *Aspergillus niger*. *J. Gen. Microbiol.* 137:2017-2023 (1991).

183. **Yamamoto, K., H. Yoshida, K. Kokame, R. J. Kaufman, and K. Mori**. Differential contributions of ATF6 and XBP1 to the activation of endoplasmic reticulum stress-responsive cis-acting elements ERSE, UPRE and ERSE-II. *J. Biochem.* 136:343-350 (2004).

184. **Ye, J., R. B. Rawson, R. Komuro, X. Chen, U. P. Dave, R. Prywes, M. S. Brown, and J. L. Goldstein**. ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. *Mol. Cell* 6:1355-1364 (2000).

185. **Yoshida, H., K. Haze, H. Yanagi, T. Yura, and K. Mori**. Identification of the *cis*-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose- regulated proteins. Involvement of basic leucine zipper transcription factors. *J. Biol. Chem.* 273:33741-33749 (1998).

186. **Yoshida, H., T. Matsui, A. Yamamoto, T. Okada, and K. Mori**. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* 107:881-891 (2001).

187. **Yoshida, H., T. Okada, K. Haze, H. Yanagi, T. Yura, M. Negishi, and K. Mori**. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the *cis*-acting element responsible for the mammalian unfolded protein response. *Mol. Cell Biol.* 20:6755-6767 (2000).

188. **Yoshida, H., T. Okada, K. Haze, H. Yanagi, T. Yura, M. Negishi, and K. Mori**. Endoplasmic reticulum stress-induced formation of transcription factor complex ERSF including NF-Y (CBF) and activating transcription factors 6alpha and 6beta that activates the mammalian unfolded protein response. *Mol. Cell Biol.* 21:1239-1248 (2001).

189. **Yoshida, H., A. Uemura, and K. Mori**. pXBP1(U), a Negative Regulator of the Unfolded Protein Response Activator pXBP1(S), Targets ATF6 but not ATF4 in Proteasome-mediated Degradation. *Cell Struct. Funct.* 34:1-10 (2009).

190. **Zapun, A., C. A. Jakob, D. Y. Thomas, and J. J. Bergeron**. Protein folding in a specialized compartment: the endoplasmic reticulum. *Structure. Fold. Des* 7:R173-R182 (1999).

191. **Zhang, K. and R. J. Kaufman**. Signaling the unfolded protein response from the endoplasmic reticulum. *J. Biol. Chem.* 279:25935-25938 (2004).

Publications

Mulder, H. J., M. Saloheimo, M. Penttila, and S. M. Madrid.
The transcription factor HACA mediates the unfolded protein response in *Aspergillus niger*, and up-regulates its own transcription.
Molecular Genetics and Genomics 271:130-140 (2004).

Mulder, H. J., I. Nikolaev, and S. M. Madrid.
HACA, the transcriptional activator of the unfolded protein response (UPR) in *Aspergillus niger*, binds to partly palindromic UPR elements of the consensus sequence 5'-CAN(G/A)NTGT/GCCT-3'.
Fungal Genetics and Biology. 43:560-572 (2006).

Mulder, H. J., I. Nikolaev.
A HacA-dependent transcriptional switch releases *hacA* mRNA from a translational block upon Endoplasmic reticulum stress.
Eukaryotic Cell. 8:665-675 (2009).

Curriculum Vitae

Harm Mulder is geboren op 9 februari 1970 te Oosterend-Texel. In 1987 behaalde hij het HAVO diploma aan de Rijks Scholen Gemeenschap te Den Burg en begon daarna aan 't Noorderhooft College in Den Helder aan de MTS opleiding petroleum en gas. In 1991 behaalde hij het MTS P&G diploma en begon dat zelfde jaar met de leraren opleiding biologie aan de Noordelijke Hogeschool te Leeuwarden, om na het behalen van de propedeuse in 1992 aan de Rijksuniversiteit Groningen te starten met de studie Biologie. In augustus 1998 studeerde hij af aan de RuG als moleculair bioloog en is als research scientist werkzaam geworden op de R&D afdeling van het Deense biotechnologie bedrijf Danisco te Kopenhagen. Van 2000 tot 2004 heeft hij bij Danisco, en binnen het Eurofung framework 5, onderzoek gedaan aan de Unfolded Protein Response van *Aspergillus niger*, wat uiteindelijk heeft geleid tot dit proefschrift. In september 2007 is hij vervolgens terug gekeerd naar Nederland en heeft zijn werkzaamheden voortgezet als scientist bij Genencor te Leiden, een divisie van Danisco.